Community-Centred Agents and Emergent Social Conventions

Bradley A. Swerdfeger
bswerd@cs.ubc.ca
University of British Columbia

ABSTRACT

In this paper, we examine techniques by which a community of autonomous agents can reach global agreement upon the use of social conventions in coordination and cooperation games. The merits of examining such convention acquisition is twofold: it can help us learn more about how humans behave in social systems, and it can explain how agents should behave in artificial social systems. We notice that one of the most established convention acquisition strategies seems to rely on some tenuous assumptions and does not represent human behaviour very well. Therefore, we invent new, community-centred convention acquisition strategies. A new environment that takes location into account is defined and these new strategies are examined against other well-established strategies in terms of convergence, stability and social welfare generation. The new strategies are concluded to be optimal in our environment and give rise to interesting human-like behaviours. Suggestions for making the simulation environment more realistic are given and future work is suggested.

1. INTRODUCTION

Agent-based simulations model a population undergoing frequent interactions with one another where we explicitly model the individual agents rather than gross statistics for the entire population[6]. Two particularly interesting types of interactions are coordination and cooperation. In these types of interactions, agents must agree on an action to receive the optimal amount of utility.

To ensure that agents cooperate or coordinate, they tend to follow *social conventions* which represent a behavioural constraint that strikes a balance between individual freedom and the goal of the agent's society. Conventions can come to exist within society in two different ways:

Designed: Laws are hardwired into agents by a designer.

Emergent: Conventions emerge organically through agents capable of learning and updating their strategy.

In this paper, we are concerned with emergent conventions, as they are the more interesting case that tends to have some practical explanatory power with regard to human behaviour.

Agents learn conventions through strategy update functions employed by the agents. For these agent-based simulations, two constraints are generally applied to the strategy update function[7]:

Semi-Locality: The strategy update function must be based on the particular agent's experiences, including only the actions taken, the actions taken by opponents, and the payoffs received. A stronger constraint called *locality* abandons the ability to refer to the actions of opponents.

Anonymity: The strategy update function cannot be based on the identities of individual agents or the names of the actions

We abandon semi-locality for some strategy update functions to allow for costless communications between agents. This seems plausible in human society since we are so well connected and can communicate over short distances with no cost other than a very short time span. Also, some well-established strategy update functions rely on cheap communication between agents[3]. However, some strategy update functions will maintain the above constraints.

In [7], Shoham and Tennenholtz define a *rational social convention*:

Rational Social Convention: Let g be a game, $V(g)$ a variable defining either the maxmin value, the set of values of the various Nash equilibria, or the set of values of the Pareto-optimal outcomes attainable in game g, and $<$ an ordering on the possible values of this variable[1]. A social convention sl is rational with respect to g and V if $V(g) < V(g_{sl})$.

Shoham and Tennenholtz also propose a simple, self-interested strategy update function, called Highest Cumulative Reward (HCR), which they claim converges on a convention that guarantees to the agent a payoff which is no less than the maxmin value of the game in coordination and cooperation games. This convention is not necessarily a Nash equilibrium, and in the case of the Prisoner's Dilemma, it is the rational convention with respect to maxmin where both agents cooperate.

There are a few assumptions that are made to guarantee this convergence to a certain probability that seem rather

[1]If the game variable refers to a set of elements we take $<$ to be an ordering over sets. With maxmin, it is straightforward.

Cite as: Title, Author(s), *Proc. of 7th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2008)*, Padgham, Parkes, Müller and Parsons (eds.), May, 12-16., 2008, Estoril, Portugal, pp. XXX-XXX.

tenuous. First of all, in their proof, they rely on pairs of agents with the same strategy playing together until other pairs of agents forget their past. What if agents don't ever forget their past? Secondly, they abandon doing empirical simulations on The Prisoner's Dilemma since it ends up being very inefficient at converging to a rational convention and therefore does not react the way that they predict (they admit as much). This fact is very unsatisfying and does not make a very good case for their claim. HCR should be evaluated transparently. Finally, they assume that the memory size is greater than the number of agents. This is not necessarily realistic, as we might want to have more agents than iterations in a simulation. Naturally, we want to exploit parallelism in our environment.

In addition to testing HCR in my own simulation environment, we would like to test new strategy update functions that are community-centred. In human society, people are generally concerned with the welfare of others, as well as themselves. They are not strictly self-interested utility maximizers. If HCR fails for some reason, then it would be good to have a strategy update function that emulates human behaviour and also converges to a rational convention that is equitable and social welfare maximizing.[2]

In [2], Richard Dawkins asserts that one of the main ideas of the theory of evolution is the fact that *genes* that are selfish, not individuals. He states that there are circumstances in which genes assure their own selfish survival by influencing individuals to behave *altruistically*. For instance, reciprocal altruism is an obvious example when there is an asymmetry in need: both agents benefit from a reciprocally altruistic interaction. This kind of 'you scratch my back, I'll scratch yours' mentality would lead agents to cooperate in the Prisoner's Dilemma.

In [4], biologist Marc Hauser describes a study where people are given challenging moral conundrums and they must decide on the most moral course of action. Despite differences in cultural or religious background, there was overwhelming agreement on the most moral outcome. This lends evidence to the fact that there is a moral sense built into our brains. The outcomes chosen were equitable and utilitarian in nature. People also chose an equitable Pareto-optimal outcome whenever possible. This suggests that the most desirable emergent convention is one that is utilitarian, but also equitable and Pareto-optimal, if possible.

These two examples of human behaviour suggest two types of community-centred strategy update functions: one that is altruistic, and one that is utilitarian.

In addition to testing HCR, we would like to compare the performance of the newly proposed, community-centred strategies against other well-established strategy update functions in terms of convergence, stability and social welfare generation in order to make sure that our new update functions are feasible replacements for more established ones.

We begin in Section 2 by definining a formal model of our simulation environment including the various strategy update functions and our evaluation methodology. In Section 3 we present the results obtained after extensive experimental simulations of communities of agents employing the various update functions. Next, in Section 4 we discuss these results in terms of performance. After that, in Section 5 we discuss interesting emergent behaviours that were noted while performing simulations and relate them to behaviours exhibited by humans. Finally, in Section 6 we conclude the paper and provide directions for future work.

[2]Note that since the games are symmetric, utilities are normalized across agents and we do not have to worry about unit conversion issues.

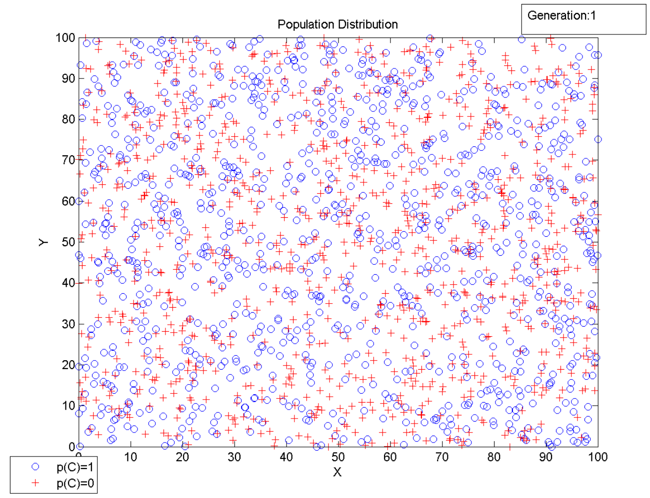


Figure 1: Example initial simulation environment.

2. THE SIMULATION ENVIRONMENT

Our simulation environment is atypical from standard agent-based simulation environments. Most simulations have a population of agents, where subsets of agents are chosen randomly to play games against each other[7][5]. However, the simulation environment detailed in this section takes location into account to model the effects of proximity. Agents wander randomly around a finite environment, playing games with proximal agents at each time-step. Agents are restricted to pure strategies, but they are able to update their strategies based on their interaction histories or communications from other agents. This section will define this environment formally.

2.1 Agents, Interactions, and Games

We begin assuming a set $N = \{1, ..., n\}$ of agents. Agents are initially assigned with a uniformly random distribution to a location, $l_i = (x, y), 0 \leq x, y \leq 100, x, y \in \mathbb{R}$ in a 100x100 space. Initially, agents are also assigned to a pure strategy $s_i \in A_i$ where A_i is the set of possible pure strategies for player i. The density of agents assigned to each strategy is specified by the experimenter. An example initial environment can be seen in Figure 1. The experimenter also declares the strategy update function for each agent, which will be explained in detail in the following subsection.

Each simulation has a number of generations. At each generation $k = \{1, ..., T\}$, an agent moves a random distance in both the x and y directions, where the direction is chosen uniform randomly in the interval $[-1, 1]$. After movement, the system uses the KD-Tree algorithm[1] to find each agent's approximate nearest neighbour. This neighbour is the agent's opponent for this generation's subsequent game-playing. If no other agent is found within a box of size 2 around an agent, then they do not play a game this generation. This *no-op* does not take any part in the strategy selection process, therefore it is ignored by the model.

Each agent has a memory M_i, which is a finite series of interactions. Each element in the agent's memory is a triple $\mu_i^k = (s_i^k, \sigma_i^k, u_i^k) \in M_i$, containing the player's move s_i^k, the opponent's move σ_i^k at time k against player i and the utility u_i^k gained from the outcome of the game. For the purposes of these experiments, each agent's memory size is set to the number of generations in the simulation, $|M_i| = 2T$, thereby giving them essentially infinite memory. Notice that the agent's memory meets the semi-local and anonymity constraints.

There are three different games that we examine in this experiment, but the agents only play one type of game in each simulation. The normal form of the games are shown in the following tables, and they are named as follows: The Prisoner's Dilemma (cooperation game), The Coordination Game, and The Coordination Game with Pareto Optimal Outcome (for short, The Pareto Game).

	C	D
C	-1,-1	-4,0
D	0,-4	-3,-3

(a) The Prisoner's Dilemma

	C	D
C	1,1	-1,-1
D	-1,-1	1,1

(b) The Coordination Game

	C	D
C	2,2	-1,-1
D	-1,-1	1,1

(c) The Pareto Game

Table 1: Games associated with this simulation environment.

Here we list the conventions that are rational with respect to the maxmin value that are desired in each game:

- In The Prisoner's Dilemma, the only desired convention is the one where both agents select 'C', or cooperate.
- In The Coordination Game, there are two possible desired conventions: where both agents select 'C' or where both agents select 'D'.
- In The Pareto Game, there is one desired convention where both agents select 'C', since this outcome is also a unique, equitable Pareto-optimal outcome.

2.2 Strategy Update Functions

The aim of this paper is to examine the effects of different strategy update functions (SUF), and the interactions between different update functions in terms of their effectiveness in producing convergence, stability and high social welfare. An SUF f_i takes a memory state or a set of communications from other agents and, based on this information, selects a pure strategy. The experimenter selects the strategy update function for each group of agents at the beginning of the simulation. The different strategy update functions are described below, and they come in two flavours: Self-interested and Community-centred.

In our simulation environment, the strategy is updated when $|M_i| > 0$ and before an agent partakes in an interaction with another agent at each generation.

2.2.1 *Self-interested Strategies*

When an agent has a self-interested strategy, they are only concerned with their own utility. Agents act selfishly and no concern is given to the community. The self-interested strategies are as follows:

Stubborn: This is the very simplest form of update function: none. Agents with this SUF will keep the pure strategy that was assigned to them for the duration of the simulation.

Cheaptalk: With this SUF, the opponent sends a costless message to the current agent indicating which action they intend to take (only one agent communicates their intentions). The agent then switches to the pure strategy that maximizes their utility against the communicated action[3].

Tit-for-tat: This strategy update function is only relevant in the Prisoner's Dilemma game. With this strategy, the agent will begin cooperating. If any agent defects against them, then they will defect in the next round then go back to cooperating. Note that agents have no identity. This means that the agent might not necessarily defect against the same agent that defected against them, which has implications that will be explained in later sections.

Highest Cumulative Reward (HCR): Proposed by Shoham and Tennenholtz[7], This local SUF is defined algorithmically as follows[6]:

1. Initialize the cumulative reward for each action to 0.
2. Play according to current action and update its cumulative reward.
3. Switch to a new action iff the total payoff obtained from that action in the latest m iterations is greater than the payoff obtained from the currently chosen action in the same time period.
4. Go to Step 2.

For our simulation, we set m to the number of games that the agent has participated in at each generation, making agents never forget their past. Formally, we can define this as:

$$f_i = \operatorname*{argmax}_{s_i \in A_i} \left(\sum_{k: s_i^k = s_i} u_i^k \right), (s_i^k, \cdot, u_i^k) \in M_i \qquad (1)$$

2.2.2 *Community-centred Strategies*

Community-centred strategies are those where consideration is given to other agents in the community. Agents might try to emulate their peers, or they might act to increase the utility of other agents the community, as well as themselves. These strategies are defined as follows:

Simple External Majority: With this semi-local SUF, agents will switch their strategy if so far they have observed more instances of it in other agents than their present strategy. Formally, it can be defined as: $f_i = \text{mode}(\sigma_i^k)$[8].

Utilitarian: Here, the agent changes their strategy such that the total utility gained by themselves and their opponent is maximized over their entire memory, therefore this SUF is semi-local. The agent chooses the action that maximizes the social welfare of the community formed by agents that he has interacted with. This can be defined formally as:

$$f_i = \underset{s_i \in A_i}{\text{argmax}} \left(\sum_{k=1}^{g} \left(u_i(s_i, \sigma_i^k) + u_o(s_i, \sigma_i^k) \right) \right) \quad (2)$$

where o is the opponent and g is the current generation.

Locally Altruistic: With this SUF, agents choose the strategy that maximizes the utility of the agents surrounding them in a certain radius, given their current strategy. This can be seen as a less local version of Cheaptalk, where the agent has no regard for their own utility, but rather the utility of their potential opponents. Defined formally:

$$f_i = \underset{s_i \in A_i}{\text{argmax}} \left(\sum_{j \in R(i,r)} u_j(s_i, s_j) \right) \quad (3)$$

where $R(i, r)$ is the set of agents within radius r of agent i.

2.3 Performance Characteristics

Walker and Wooldridge, in [8], define two statistics to assess the performance of strategy update functions: norm convergence and average number of strategy changes. I will also add social welfare as a performance characteristic to examine which SUFs are best for an entire community of agents.

Norm Convergence ($\mathcal{C}_k$): One of the aims in designing a good strategy update function is to have the agents in a system converge on a strategy as quickly as possible. Convergence is defined as the fraction of agents using the most popular strategy at generation k:

$$\mathcal{C}_k = \frac{\max_{s_i \in A_i} |ch(s_i, k)|}{|N|} \quad (4)$$

where $ch(s_i, k)$ is the set of agents that have chosen pure strategy s_i at time k.

Average Strategy Changes ($\mathcal{N}_k$): We also want to design SUFs that are *stable*: meaning that it is disadvantageous to keep switching from one strategy to another. Intuitively, in real-world situations, switching strategies might incur a cost. For example, if a company decides to switch from one operating system to another. We want to minimize the average number of changes made by an agent at each generation k:

$$\mathcal{N}_k = \frac{\sum_{i \in N} sc(i, k)}{|N|} \quad (5)$$

where $sc(i, k)$ returns 1 if agent i changes strategies at time k, 0 otherwise.

Social Welfare ($\mathcal{W}_k$): Since we are concerned with coordination games and the prisoner's dilemma, we want the adapted convention to be good for every agent in the community. We want all agents in the community to be happy, rather than a select few individual agents. For this reason, we record the average utility of each agent at generation k:

$$\mathcal{W}_k = \frac{\sum_{\gamma \in \Gamma_k} \sum_{i=1}^{\eta} u_i(\gamma, s_i, s_{-i})}{\eta\, |\Gamma_k|} \quad (6)$$

where Γ_k is the set of games played in generation k, η is the number of agents that play in an instance of a game (in this case, $\eta = 2$) and $u_i(\gamma, s_i, s_{-i})$ is the utility gained by player i in game γ given that they played strategy s_i and their opponents played strategies s_{-i}.

	Coord.	Pareto	Prisoner's
Cheaptalk	90	88	100
Tit-for-Tat	-	-	90
HCR	79	96	61
Majority	55	57	-
Utilitarian	94	96	100
Altruistic	74	100	100

Table 2: Experimentally Determined Values of $\mathcal{C}_k$, for $k = 250$.

3. EXPERIMENTAL RESULTS

The six strategy update functions were tested on each of the different games. Each simulation ran for 250 generations and results are averaged over 25 iterations to avoid any statistical abnormalities. Each experiment consisted of 2000 agents ($N = 2000$) with half of the agents initialized to the pure strategy $p(C) = 1$ while the other half were initialized to $p(C) = 0$. For the following results, all agents were given the same strategy update function. Interactions between SUFs will be examined in section 5. For the Locally Altruistic SUF, the radius is set to 5.

3.1 Convergence - $\mathcal{C}_k$

The experimentally determined values for $\mathcal{C}_k$, for $k = 250$, are given in Table 2. Note that these values are displayed as percentages.

Based on the results in Table 2, we make the following observations:

- The SUFs with $\mathcal{C}_k = 100$ convergence actually converge to 100% at about $k \approx 50$, therefore these strategy update functions are exceptionally fast at converging.

- The Majority update function is useless with equal numbers of uniform randomly distributed agents. The only thing contributing to any convergence is randomness.

- HCR is fairly slow at converging in The Coordination Game and The Prisoner's Dilemma. This is due to the fact that we have uniform randomly distributed agents, therefore on average, agents will play against half that cooperate and half that defect until the population becomes close to reaching a convention.

- The Altruistic update function does not converge very well for the coordination game. This is due to the fact that there is no unique rational convention to converge to, therefore subsets of co-located agents choose the

	Coord.	Pareto	Prisoner's
Cheaptalk	10.5	11.4	0.0
Tit-for-Tat	-	-	9.8
HCR	1.0	0.3	332.2
Majority	2.0	1.7	-
Utilitarian	0.4	0.4	0.0
Altruistic	1.3	0.0	0.0

Table 3: Experimentally Determined Values of $\mathcal{N}_k$, for $k = 250$. All values are multiplied by 10^3.

	Coord.	Pareto	Prisoner's
Cheaptalk	0.96	1.06	-3.00
Tit-for-Tat	-	-	-2.81
HCR	0.65	1.86	-2.22
Majority	0.46	0.87	-
Utilitarian	0.86	1.85	-1.00
Altruistic	0.75	2.00	-1.00

Table 4: Experimentally Determined Values of $\mathcal{W}_k$, for $k = 250$.

same strategy, but distal agents may not. This community building phenomenon will be elaborated upon in section 5.1.

- The non-HCR self-interested SUFs converge rather steadily in general, and HCR converges quite will in The Pareto Game.
- The new Utilitarian and Altriustic SUFs compete very well with the other strategy update functions; especially when there is a unique rational convention.

3.2 Strategy Changes - $\mathcal{N}_k$

One possible issue with the Cheaptalk and Tit-for-Tat functions is the number of strategy changes that an agent might make. If these changes are costly, then the nice convergence properties could be rendered completely moot. Let us examine the experimentally determined values for $\mathcal{N}_k$ for $k = 250$ in Table 3. We can make the following observations from Table 3:

- Tit-for-Tat is not very stable and Cheaptalk is not stable for the coordination games. This is not surprising as Tit-for-Tat is a very dynamic strategy update function since it makes at least one strategy change for every defection. Cheaptalk is also dynamic since it does not rely on the agent's memory or gross statistics based on surrounding agents.
- HCR is exceptionally unstable in The Prisoner's Dilemma. This is consistent to the findings by Shoham and Tennenholtz[7].
- The Majority SUF is surprisingly stable. This is likely due to the fact that the strategy relies on the memory of the agent and at $k = 250$, the majority strategy likely does not change very quickly since the memory is so well-established. Each additional generation has $1/k$ weight.
- The Altruistic and Utilitarian SUFs are very stable, especially for games with a unique rational convention.

3.3 Social Welfare - $\mathcal{W}_k$

Now we want to look at how much utility agents get from converging onto a social convention. If agents converge on an irrational convention, they may be suffering while they could be converging just as quickly on a better convention for themselves and everyone else. For this reason, we observe the experimentally determined values for $\mathcal{W}_k$ for $k = 250$ in Table 4: Note that the utility in The Coordination Game ranges from $[-1, 1]$, The Pareto Game ranges from $[-1, 2]$ and The Prisoner's Dilemma ranges from $[-4, -1]$. We make the following observations from Table 4:

- Cheaptalk and Tit-for-Tat converge to the Nash equilibrium in The Prisoner's Dilemma.
- HCR converges to the rational convention in The Pareto Game, while Cheaptalk appears to be converging to the non-rational Nash equilibrium. HCR converges to a single rational convention in The Coordination Game, but very slowly.
- The Utilitarian and Altruistic SUFs, and eventually HCR always converge to a unique rational convention.

4. DISCUSSION OF PERFORMANCE

In this section, we will discuss the performance of the various strategy update functions. We will begin with discussing convergence, and then discuss stability and social welfare, paying most attention to the functions that perform well in earlier areas, but fail to perform in subsequent measures.

Walker and Wooldridge use only majority-based strategy update functions to observe emergence in agent-based simulations[8]. In Table 2, it is apparent that majority-based SUFs will not perform well in our experimental environment. This is due to the fact that we have an equal number of agents equally distributed in space. Any convergence to a single strategy will be due to an unevenness caused by the generation of random numbers; not by the virtue of the strategy update function at selecting a proper strategy. Furthermore, majority-based strategy update functions do not take utility into account; just because something is popular, does not mean that it is good for you.

It was observed that the Cheaptalk, Utilitarian and Altruistic strategy update functions converged extremely quickly in the Prisoner's Dilemma. Despite the high numbers for the other convergence statistics, they do not compare to these SUFs since they converged at $k \approx 50$ rather than being still in the act of converging at $k = 250$. It should also be noted that all of these functions were comparable in the coordination games in terms of convergence, with the exception of the Altruistic function, which was slow to converge in The Coordination Game, but exceptionally fast in The Pareto Game. This makes sense because the Altruistic function maximizes utility in an area around the agent. If there is an equal number of agents following each Nash equilibrium, then the entire area will quickly converge to the Pareto-optimal solution as this yields greater social welfare. However, the interference of the boundaries of co-located communities observable in Figure 2a causes this strategy to converge slowly in The Coordination Game since no strategy is obviously better.

Similarly to the results found by Shoham and Tennenholtz, HCR is extremely slow to converge to a rational convention in The Prisoner's Dilemma[7]. We also found that it was quite slow to converge to a convention in The Coordination Game, when there are multiple rational conventions. This is contrary to their findings, but we have changed the simulation environment. However, HCR does manage to converge quickly in the Pareto Game.

Now that we have established that every strategy update function other HCR and Majority converges rather well for all cases, we can see how they stack up in terms of stability.

The only really notably unstable SUFs, other than the obvious HCR, are Cheaptalk in the coordination games and Tit-for-Tat. The instability of Cheaptalk seems obvious in the coordination games since the agent that receives the communication will always change their strategy to that of their opponent, whether it is Pareto-optimal or not. With equally distributed agents following different strategies, these strategy changes will be very frequent. Tit-for-Tat will cause two changes in strategy for every defection followed by a cooperation, therefore this strategy is not ideal when changing strategies incurs a cost.

At this point, the only SUFs that really survive scrutiny for this simulation environment are Utilitarianism and Altruism.

Not only do the Utilitarian and Altruistic strategy update functions converge quickly and remain stable, but they also converge to the desired rational conventions. For this reason, the Altruistic and Utilitarian strategy update functions seem more promising than any of the self-interested strategy update functions.

All performance characteristics considered, the Altruistic SUF seems to perform the best in a coordination game where there is a Pareto-optimal equilibrium and communication is costless, and the Utilitarian update function seems to perform best for The Prisoner's Dilemma and Coordination Game.

If locality is desired, then the Utilitarian strategy update function is an excellent replacement for HCR. We can still converge to the rational convention while only sacrificing locality for semi-locality. This is a very plausible and inexpensive constraint relaxation.

5. INTERESTING EMERGENT BEHAVIOURS

There are some notable behaviours that occur in some of the simulations above, as well as when interactions between different strategy update functions are introduced. These interesting behaviours include: community building, bullying, and enlightenment. As one can see from the names of the behaviours, they will have interesting parallels and implications on the way that people behave in real society.

5.1 Community Building

A community can be defined as a set of co-located individuals that share similar customs, values, and conventions. Groups of individuals that are co-located tend to develop similar conventions because there is a need within society to cooperate in order to succeed individually, and as a whole. A trivial example of this is with greeting customs: a handshake is a completely arbitrary way to greet someone, but it serves to let another person know that you wish to act in a civil manner. However, the convergence on a particular convention occurs mainly because agents are co-located and must interact with one another on a regular basis. A separate community that is distal in space might develop a completely different convention that serves the same purpose. For instance, in some Eastern countries they bow to greet someone instead of performing a handshake. These arbitrary acts perform the same purpose, but were developed separately by isolated communities.

In Figure 2, we can observe that this form of co-location-based community building and convention adaptation is apparent with the Altruistic and, to a lesser extent, Utilitarian strategy update functions. The agents eventually conform to a single convention, but if we isolated them, we would observe stability within both groups of agents. The effect is so strong that it even works, to a degree, in a game with a Pareto-optimal outcome. This phenomenon shows that the Utilitarian and Altruistic strategy update functions might be good models for how people come to converge on a completely arbitrary convention that may or may not award less utility than another arbitrary convention.

5.2 Enlightenment

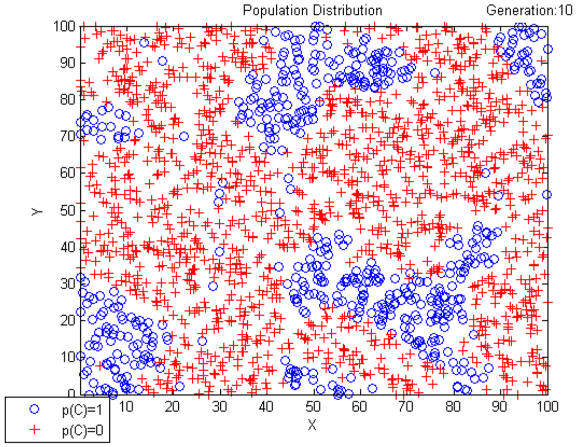
The fact that agents have agreed to conform to a specific convention does not mean that it is the best convention possible. Co-located agents might be ignorant to better ways, happily conforming to a Pareto-dominated equilibrium when they could be converging to a Pareto-optimal equilibrium. For instance, a community might believe that horses are merely a good source of food, therefore they hunt them. A distal tribe might also have horses nearby, but they have domesticated them and treat them with respect as contributing members of society. The horses and humans have formed a symbiotic relationship of labour for nutritious food and care. What happens if a member of the domestication group goes to the group that hunts horses for food and teaches them how to domesticate the horses so that they can increase their labour capacity? If the members of the horse-hunting tribe see the potential gains in utility and are not too stubborn to adopt a new way of life, they will increase the social welfare of the group (and the horses) by adopting domestication.

In Figure 3, we set the Pareto-optimal outcome to be 5 times better than the Pareto-dominated equilibrium. The agents that follow the Pareto-optimal convention are given a Stubborn SUF and their density is set to 0.1. The agents following the other equilibrium are given a Utilitarian SUF and their density is set to 5. Notice how the population slowly converges to the Pareto-optimal outcome. Given a finite memory or periodic memory resets, this convergence would occur with much greater speed (corresponding to agents in the population dying out over time and passing their teachings to their children). Take note as well that this convergence to the Pareto-optimal convention would also be attained by using the HCR strategy update function. This phenomenon demonstrates how a utility maximizing strategy update function coupled with a memory will have the capacity to be 'enlightened' with a strategy that gives a higher amount of utility.

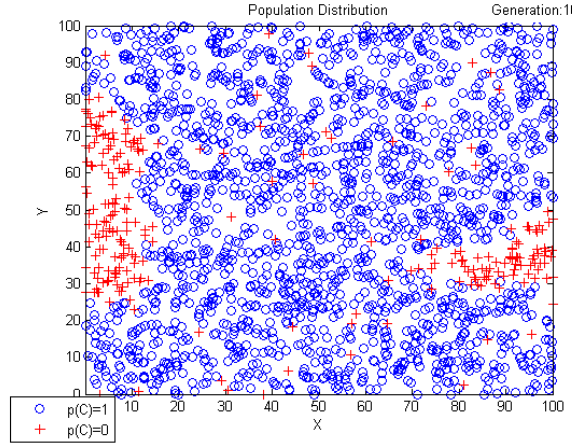
5.3 Bullying

What would happen if instead of being enlightened, a cooperating population is invaded by a few *bullies* that maximize their own utility by defecting and taking advantage of the polite, cooperating population? This might occur in an anonymous online auction system, where a small number of the sellers are dishonest and accept the money, but fail to send the item to the winning bidder.

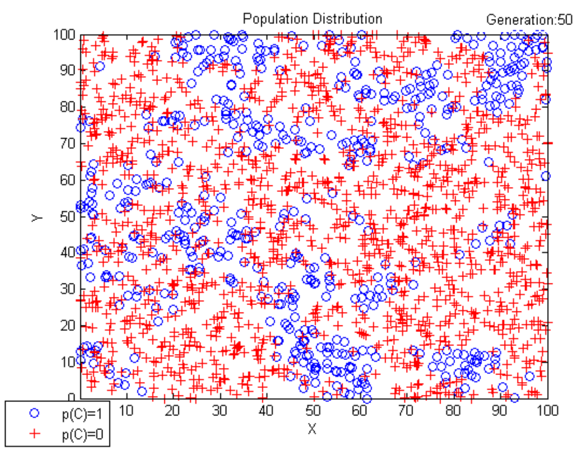
Figure 4 demonstrates something similar to this situation



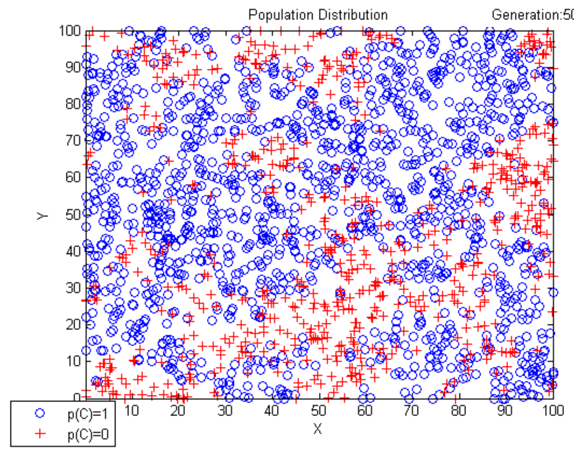
(a) Locally Altruistic agents playing The Coordination Game at $k = 10$.



(b) Locally Altruistic agents playing The Pareto Game at $k = 10$.

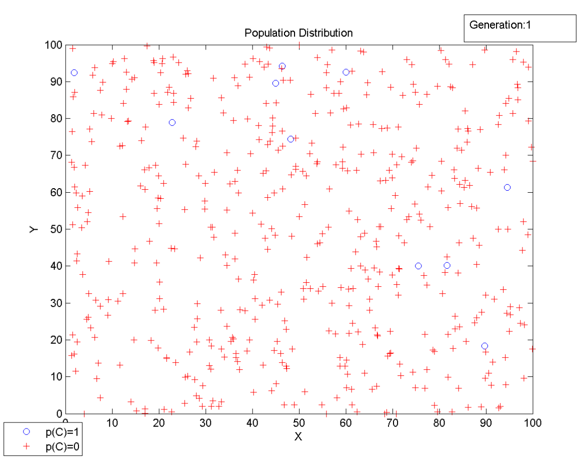


(c) Utilitarian agents playing The Coordination Game at $k = 50$.

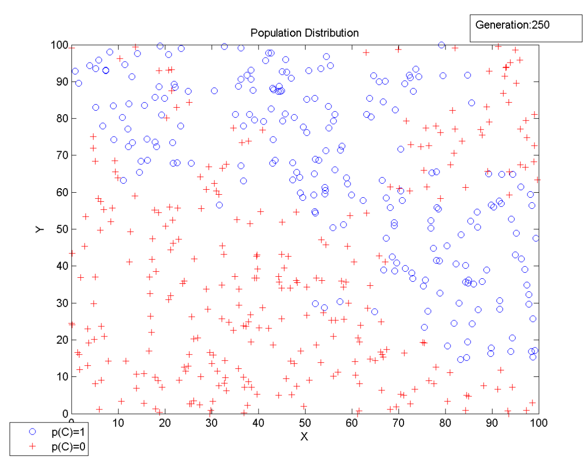


(d) Utilitarian agents playing The Pareto Game at $k = 50$.

Figure 2: Community-centred agents playing coordination games and forming communities.

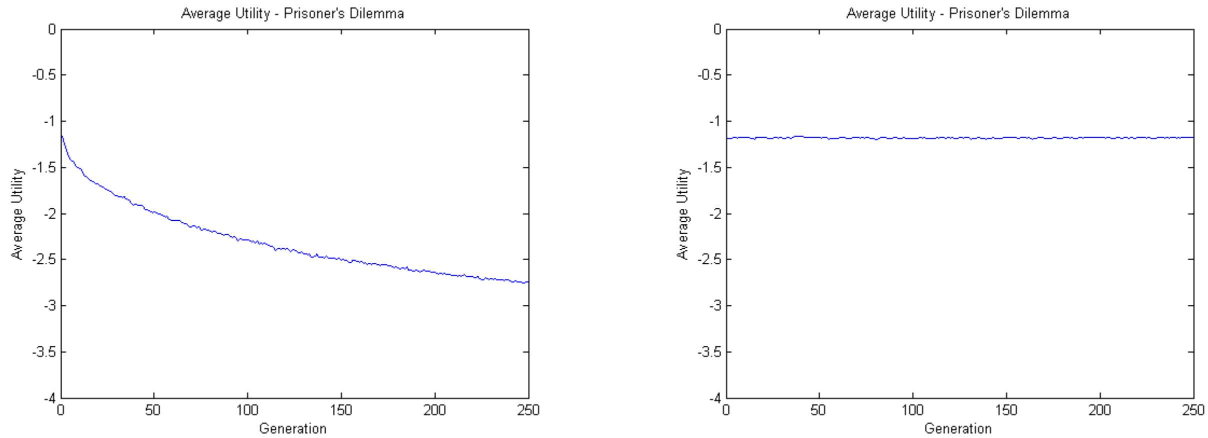


(a) The population distribution at $k = 1$.



(b) The population distribution at $k = 250$.

Figure 3: Utilitarian agents interacting with stubborn, enlightened agents following a rational convention.



(a) Average utility over time using the Tit-for-Tat strategy. (b) Average utility over time using the Utilitarian function.

Figure 4: Cooperating agents interacting with less-numerous stubborn, defecting agents in The Prisoner's Dilemma.

where we use the Prisoner's Dilemma and a small number of stubborn, defecting agents. The density of the defecting agents is set to 1, while the density of the cooperating agents is set to 10. Figure 4a shows the average utility of the agents if the cooperating agents use the Tit-for-Tat strategy update function. Figure 4b shows the average utility of the agents if the cooperating agents use the Utilitarian SUF. As you can see, the Tit-for-Tat strategy results in a vicious cycle of defection where the population eventually converges to a convention of defection (which is the Nash equilibrium). However, if agents adopt a Utilitarian strategy, they lose against the defecting agents, but the cooperating agents do not end up defecting and obtain a social welfare close to that where all agents follow the rational convention. Unfortunately, the small population of bullies win out the most, but that is a small price to pay for greater social and personal welfare.

The obvious solution to this problem is to add identity, and thereby, reputation to the agents. In this case if a Tit-for-Tat strategy was adopted where the defection was only against specific agents that defected in the past, the Tit-for-Tat strategy would perform better than the Utilitarian function. Bullies would be punished, and the rational convention would be maintained between cooperating agents. The addition of identity and reputation to agents is used by eBay and discourages honest users from partaking in transactions with users that have a poor reputation.

6. CONCLUSIONS AND FUTURE WORK

This research was motivated by a concern that the Highest Cumulative Reward strategy update function might be slow or unable to converge to a rational social convention. We wanted to find strategy update functions that would be able to always converge to a rational convention that fits human moral standards and that would be realistic in a human social setting. Furthermore, we wanted to assess the performance of various popular strategy update functions in this new environment to see which perform well in terms of convergence, stability and welfare.

We showed that the Utilitarian, and Altruistic strategy update functions perform best in terms of convergence and stability. Furthermore, we showed that the Utilitarian and Altruistic SUFs usually[3] converge to the rational convention, as they are community-centred and concerned with the welfare of other agents in the community. HCR, on the other hand, converges to the rational convention, but does so exceptionally slowly compared to the new strategy update functions.

In general, utility-concerned, community-centred strategy update functions have desirable properties for coordination and cooperation games in settings with anonymous agents.

One important assumption that was made was the fact that the game being played by the agents was well-known. Future work could examine if community-centred agents perform well in stochastic or Bayesian environments.

In addition, the distribution of the agents' location and movement was extremely simplistic. Future work could also examine if the results found in this simulation would extend to more complex movements and agent distributions and what kind of interesting, human-like behaviours emerge.

Subsequent work could also examine what kind of new strategy update functions perform well for coordination and cooperation games, other than Tit-for-Tat when agents are given identity. Social structure could also be added to the environment, where, for instance, some agents are authorities.

In fact, any relaxation of assumptions that could make the simulations more world-like are desired and should be examined with respect to emergent conventions in order to better understand how humans can act optimally, rationally, and morally in a social and individual welfare maximizing fashion.

The ability to improve and extend this research is boundless and should be tackled with vigour.

7. REFERENCES

[1] J. L. Bentley. K-d trees for semidynamic point sets. In *SCG '90: Proceedings of the sixth annual symposium on Computational geometry*, pages 187–197, New York, NY, USA, 1990. ACM.

[3]The 'always' part would have to be proved mathematically, which is left for future work

[2] R. Dawkins. *The God Delusion*. First Mariner Books, 2005.

[3] J. Farrell and M. Rabin. Cheap talk. *The Journal of Economic Perspectives*, 10(3):103–118, 1996.

[4] M. Hauser and P. Singer. Morality without religion. *Free Inquiry*, 26(1):18–19, 2006.

[5] J. Oechssler. Competition among conventions. *Computational & Mathematical Organization Theory*, 5(1):31–44, 1999.

[6] Y. Shoham and K. Leyton-Brown. *Multi-Agent Systems*. Cambridge University Press, 2008.

[7] Y. Shoham and M. Tennenholtz. On the emergence of social conventions: modeling, analysis, and simulations. *Artificial Intelligence*, 94(1):139–166, July 1997.

[8] A. Walker and M. Wooldridge. Understanding the emergence of conventions in multi-agent systems. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 384–389, Menlo Park, CA, 1995. AAAI Press.