Agents with Moral Sentiments in an Iterated Prisoner's Dilemma Exercise

Ana L. C. Bazzan
Department of Computer Science
University of Massachusetts
bazzan@cs.umass.edu

Rafael H. Bordini and John A. Campbell
Department of Computer Science
University College London
{R.Bordini, J.Campbell}@cs.ucl.ac.uk

Abstract

We present the results of a simulation of the Iterated Prisoner's Dilemma where some agents behave rationally and others have moral sentiments towards those from the same social group. They suggest that in societies where agents can possess such characteristics, to behave rationally, in the usual sense in Game Theory, is not the best attitude for good performance in the long run, both individually and for the group.

Introduction

Understanding the ability to show cooperative behaviour has concerned scientists in many areas of human knowledge. The Prisoner's Dilemma (PD) has been used as a metaphor to formalise the conflict between two individuals who can involve themselves in either mutual support (cooperation) or selfish exploitation (defection). Since there is a risk for one of the individuals to end up with the *sucker*'s payoff, to defect is always a good tactic to follow.

However, when played repeatedly, in which case it is called the Iterated Prisoner's Dilemma (IPD), mutual defection is no longer the only solution, although it is the rational one. This was verified in Axelrod's computer tournament (1984), in which a program representing the tactics called Tit-For-Tat (TFT) won by cooperating initially, then repeating what the opponent did in the last round.

These ideas have been employed in the field of Multi-Agent Systems in order to explain the achievement of cooperation and coordination. However, little work has been devoted to the IPD for social agents, particularly agents with emotions. This may mirror the fact that there is actually little work on cooperative behaviour among primates, let alone human beings.

One reason for this may be the failure of theories based on rational choice to account for *social* action and choice as claimed by Conte & Castelfranchi (1995). Social agents are constantly involved in planning their actions, and they need to reason about sub-goals that arise in this process. Game theory tends to treat these goals in a process of choice for each agent in isolation and exclusively from its own point of view. For instance, a goal does not lead to any attempt to modify any mental states of an opponent.

Below, we examine various issues of relations among social agents in an IPD context, including moral and philosophical aspects such as why people are able to keep their promises once they agree to cooperate and why people behave altruistically. The ideas on moral sentiments that inspired this work are presented next. We then describe a simulation that assesses those ideas, give the results and some analysis, and finally mention possible extensions to the work.

Moral Sentiments

In a recent (1996) publication on the "Origins of Virtue", Ridley makes the point that *Moral Sentiments* (emotions like generosity towards others and guilt for not having played fair with someone) prevent us from being *Rational Fools* (see Chapter 7). Rational fools act to maximise their gain in the short term, which does not pay off in the long run because people do not reciprocate with those who have proven selfish in the past. Moral sentiments lead us to sacrifice rational decisions, yet they are of fundamental importance to social relations inasmuch as they allow us to create a reputation as altruistic people. Altruism, which most people praise as a virtue, will lead an altruistic person to experience the generosity of others as well, when it is needed. However, these same emotions drive us to want those who belong to the same social group to be somewhat self-interested, which is better for the group too. We are particularly altruistic with people from the same social group or who share our genes.

In other words, moral sentiments are decisive in the dilemma between getting the maximum out of an opportunity or being cautions about the future. They are part of our highly social nature, to favour our genes' long-term advantage. They are also a guarantee of our commitments, which makes complex social relations possible; and that too stands in our long-term advantage. Morality can be seen as a set of instincts preventing us from being selfish, which is for our own sake in the long run. When, however, people do some-

thing that is not rational and does not pay off even in the long run (true altruism), they are falling prey to sentiments originally designed to obtain other people's trust, which is convenient for real life's "prisoner's dilemmas"[1].

These are the general ideas that have inspired the conception of the simulation we describe below. Similar ideas were also developed by Simon (1990).

Agents with Moral Sentiments and IPD

We are interested in analysing the performance of groups of agents playing the IPD in the presence of agents with moral sentiments (*altruistic agents*) or rational fools (*egoistic agents*) in that group; we also analyse the individual performance of each type of agent in the possible contexts (in a group where all agents are exclusively of one of the two types just introduced or in a mixed group). The fact that two agents belong to the same group in this simulation implies that they have some sort of emotional liaison (that is, they may belong to the same family or social group). Agents should be interested in the good performance of their group as a whole, as well as their own, since the social group provides also a base for support in case the agent itself is not performing well. However, that is not true for the egoist agents, which always seek to obtain maximum points, no matter with whom they are playing. We aim to demonstrate Ridley's point that the presence of agents with moral sentiments favours the well-being of the whole social group.

In this experiment, the pairs of agents playing the IPD are chosen randomly from all groups. Agents pay one point to play (representing the effort one puts into interacting socially). If an agent runs too low in points (because, e.g., it has played with several egoists) it is said to be *bankrupt*, in which case it is not allowed to play, but it recovers by earning one point so that it can afford to pay for the play in the next step of simulation. The points agents earn by playing are the standard amounts for the PD payoff matrix: $R = 3$, $S = 0$, $T = 5$, and $P = 1$. In order to determine the *wealth state* of an altruistic agent (which influences how it will play), we compute the average number of points through the completed simulation steps. According to certain thresholds on this average, an agent's state can be *wealthy*, *medium*, or *poor*.

Egoistic agents defect in all interactions (ALLD). Altruistic agents play Tit-For-Tat (TFT) when playing with agents from other groups, so they play a fair game here. When interacting with agents from the same social group, an altruistic agent plays with Moral Sentiments (MS), which is explained next. If the altruist is in a wealthy state and the opponent (from the same group) is in a poor one, the agent cooperates despite the fact that it knows that the opponent will defect; it does this to help the opponent to earn more points (remember that the agent has a special feeling for this opponent because they are both from the same social group). A poor altruist cooperates if the opponent is in either a medium or a poor state, because the poor altruist does not try to recover by taking advantage of a peer that is not in a wealthy state. However, an altruistic agent will defect when in a poor state if the opponent is in a wealthy one and belongs to the same group. This is a social mechanism to allow agents to recover from a bad state, which is possible through the altruism of those from the same group. Briefly, an altruist cooperates with those of the same group unless it is in a poor state and the opponent is wealthy.

Next, we analyse the results of a particular configuration of the simulation whose main features we have just described.

"To Be Good Pays Off"

We consider the particular results for the simulation of a society of agents composed of 3 groups, each one with 4 agents. One of the groups has no egoist agents (G1), another is a mixed group (G2) with 2 egoists and 2 altruists, and the third group (G3) is composed solely of egoists. Figure 1 shows the performance of each group for the average of 50 repetitions of the simulation.

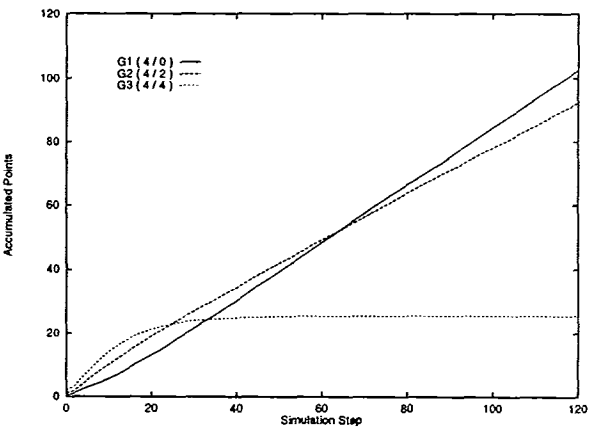

Figure 1: Simulation with 3 Groups and 12 Agents, of which 6 are Egoists.

The results show clearly that, in the long run, homogeneous groups of altruistic agents accumulate more points than any other type of group. The altruists are not rational fools; they compromise their present possibilities of gain to make sure they will do well in the future. The whole group performs well because individual failures are compensated by the generosity of those doing well, avoiding bankruptcy. Therefore, to be altruistic does pay off in the long run! We have also verified that homogeneous groups of egoists perform very well only in the short term. Their selfishness in the game compromises their reputation: once the

[1] Recall that, as Ridley puts it, this is only a dilemma if one does not know if one can trust one's accomplice.

agents in a society have found out about their character, they suffer retaliation (characteristic of the TFT). At this point, egoists only earn enough points to survive, that is, pay to play in the next step (which is why the graph for G3 in Figure 1 is horizontal, once it stabilises). Briefly, their performance has a logarithmic curve. Mixed groups have an intermediate performance, but they clearly do not exhibit the catastrophic effect of a group where there are no altruists. The presence of some altruists there assures the relative development of the group.

We have also inspected the performance of individual agents, e.g. as in Figure 2 (again 50 repetitions). There, we use four different types of line, for the different types of agents and group contexts[2]: altruistic agents in a homogeneous group (AH), altruists in a mixed group (AM), egoists in a mixed group (EM), and egoists in a homogeneous group (EH).

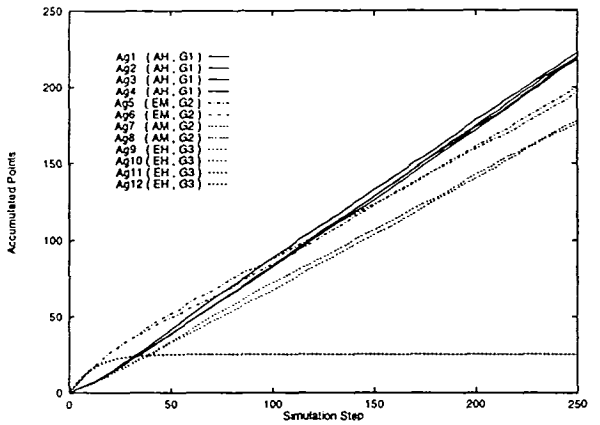

Figure 2: Individual Performances for the Simulation with 3 Groups and 12 Agents, of which 6 are Egoists.

This gives us further insights. Unexpectedly, egoists from the mixed group (EM) perform better than the altruists (AM) in that group. This is because these egoists can count on the generosity of the altruists in their group, who cooperate with them despite the fact that they are selfish, for the sake of kinship welfare. However, we notice that the presence of the egoists is harmful for all members of the group in the long run, since they would all be performing better if the egoists would stop being rational fools (see the performance of the agents in group G1).

From other configurations of the simulation (e.g., different numbers of groups, agents and percentage of each type of agent in the groups), whose graphs we do not have the space to show here, we have also reached further conclusions:

- the more egoist agents (keeping the same configuration), the worse the general performance of the society (as a whole); that is, the fewer total points (from all agents in all groups) are accumulated in the same period;
- the more egoists in a group, the faster the group collects points initially, but the worse its eventual performance will be after some time; the fewer egoists in a group, the better the group's performance;
- the smaller the percentage of egoists in a mixed group, the better the performance of each individual egoist agent (some interesting quantitative details remain to be investigated).

Several other ideas can be exploited by changing the various parameters to the simulation. We intend to report on these in further papers.

Conclusions

Our results suggest strongly that rational fools maximise their earnings in the short term but compromise their performance in the long run. The results also clearly show that the more altruists in a group, the better they, and the group as a whole, perform. Accordingly, in a society where agents have emotions, to behave rationally (in the classical sense in Game Theory) may not be the best attitude in the long run.

Other configurations of agents and groups of agents were considered and will be presented in a future paper. We shall also present there the results for a slightly different version of this simulation, where agents are not given points to recover but are allowed to receive negative points. Besides, in that version, not all agents play at every simulation step; the more agents are allowed to play in each step, the wealthier the society, since it denotes that more opportunities for business (in the form of PD interaction) exist at each instant in the society. Finally, we also plan to introduce some sort of learning process from which agents can discover the characters of others and eventually refuse them as partners in future social interactions, as Ridley argues to be the case among humans.

Acknowledgements

Rafael H. Bordini gratefully acknowledges the Brazilian agency CAPES for the grant for his PhD at UCL.

References

Axelrod, R. 1984. *The Evolution of Cooperation.* New York: Basic Books.

Conte, R., and Castelfranchi, C. 1995. *Cognitive and Social Action.* London: UCL Press. 215p.

Ridley, M. 1996. *The Origins of Virtue.* International Series in Computer Science. London: Viking Press.

Simon, H. A. 1990. A mechanism for social selection and successful altruism. *Science* 250(4988):1665–1668.

[2] Altruists and egoists have different performances when they are in homogeneous or mixed groups, which is why the distinction is important.