Stable cooperation in changing environments

(Short Paper)

Colm O'Riordan
Dept. of Information Technology
NUI, Galway
Ireland.
colm.oriordan@nuigalway.ie

Humphrey Sorensen
Dept. of Computer Science
University College Cork
Ireland.
h.sorensen@cs.ucc.ie

ABSTRACT

This paper addresses the issue of emergence of robust cooperation among self-interested agents interacting in N-player social dilemma games. A series of graphs are created each exhibiting a different level of community structure; we show the influence that community structure has on the emergence of cooperation. A strategy set that represent a form of generalised tit-for-tat is used. Two forms of uncertainty in the environment are also modelled. The influence of the game length is also explored.

Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]: Multi-agent systems

General Terms

Experimentation

Keywords

Cooperation, spatial organisations, uncertainty

1. INTRODUCTION

Social dilemma games such as the prisoner's dilemma[1] and its variants have been studied to model the interactions between autonomous agents in multi-agent systems. In this paper we adopt a generalised N-player version of the prisoner's dilemma based on the formalism of Boyd and Richerson [2]. In this game, N players simultaneously interact, either cooperatively or not.

The strategies used in this paper are generalised versions of *tit-for-tat*. The *tit-for-tat* strategy [1] and can be described as follows: cooperate on the first move and then mirror the opponents' behaviour. This can be generalised to work in the N-player iterated prisoners dilemma. Instead of simply mirroring an opponent's behaviour, these strategies cooperate if a sufficiently high number of cooperators existed in the previous iteration.

In this paper, we explore the emergence of cooperation among agents participating in an N-player social dilemma; agents are arranged and their interactions constrained by a graph topology exhibiting a level of community structure. We expand on previous work to show that with simple learning mechanisms cooperation emerges which is robust in the presence of high levels of noise. Furthermore, we show that with the presence of small levels of noise the society of agents can adapt to dramatic changes in the environment. The influence of game length is also explored.

Cite as: Stable Cooperation in Changing Environments (Short Paper), Colm O'Riordan, Humphrey Sorensen, *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. 1529-1532.

2. RELATED WORK

N-player dilemmas are characterised by having many participants, each of whom may choose to cooperate or defect. Any benefit or payoff is received by all participants; any cost is borne by the cooperators only.

Defection represents a dominant strategy, i.e. for any individual, moving from cooperation to defection is beneficial for that player. However, if all participants adopt this dominant strategy, the resulting scenario is sub-optimal. If any player changes from defection to cooperation, the performance of the society improves, i.e. a society with $i + 1$ cooperators attains a greater payoff than a society with i cooperators.

Previous work adopting evolutionary simulations has shown that without placing specific constraints on the interactions, the number of participants or the strategies involved, the resulting outcome is that of defection[8][12].

In studying the two-player game, many researchers have explored the effect of placing spatial constraints on the population of interacting agents. These include, among others, experimentation with grid size and topology [7], small world[11] and scale-free graphs[9].

In this paper, we are interested in one key property of a graph: that of *community structure*. This property has also been explored in recent work[4]. A graph is said to have a community structure if collections of nodes are joined together in tightly knit groups between which there are only looser connections. This property has been shown to exist in many real-world social networks[6].

Several researchers have addressed related issues to changes in the environment: evolutionary models and co-evolutionary models where the population is changing over time and hence too is the fitness landscape, studies in the viscosity of populations[5], changing or reversing the environment [3] or by introducing noise into the model whereby agents actions are mis-implemented or mis-interpreted by other agents[10].

3. MODEL

3.1 Graph structure

In the simulations described in this paper, agents are located on nodes of an undirected weighted graph. The weight associated with any edge between nodes represents the strength of the connection (and the likelihood of these agents participating together in games) between the two agents located at the nodes. We use a regular graph of degree four neighbours. We use two different edge weight values in each graph: one (usually, a higher value) associated with the edges within a community and another associated with the edges joining agents in adjacent communities. Both weights used are in the range [0,1]. The graph is depicted in Fig. 1, where the thicker lines represent intra-community links (larger value as edge weight) and the thinner lines indicate inter-community links between neighbouring communities. The rectangles of thicker lines represent communities.

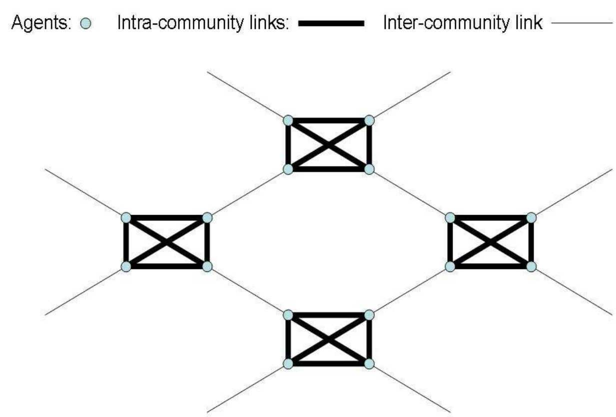

Figure 1: Graph structure

3.2 Agent Interactions

Agent stratgies are defined by two parameters—first move and a threshold. The first move specifies what to do on the first turn in an interaction; the second parameter is used to calculate a threshold, τ (represented as a value in the range 0 to 1). If there are less cooperators on turn i than τ times the number of particpants, then the agent defects on turn $i+1$. Othewise, the agent cooperates.

Agents interact with their neighbours in a N-player prisoner's dilemma. The payoffs received by the agents are calculated according to the formula proposed by Boyd and Richerson [2], i.e. cooperators receive $Bi/N - c$ and defectors receive Bi/N, where B is a constant (in this paper, B is set to 5), i is the number of cooperators, N is the number of participants and c is another constant (in this paper, c is set to 3).

Each agent may participate in several games. The algorithm proceeds as follows: for each agent a in the population, agents are selected from the immediate neighbourhood of agent a to participate in the game with a probability equal to the edge of the weight between the nodes. An agent's fitness is calculated as the average payoff received in the interactions during a generation.

3.3 Learning

Agents may change their behaviours by comparing their payoff to that of neighbouring agents. We adopt a simple update rule whereby an agent, following each round of games, updates their strategy to that used by more successful strategies. These neighbours are chosen according to the weight of the edge between agent and neighbour. Let $s\_adj(x)$ denote the immediate neighbours of agents x chosen stochastically according to edge weight. The probability of an agent x updating their strategy to be that of a neighbouring agent y is given by:

$$\frac{w(x,y).f(y)}{\Sigma_{z \in s\_adj(x)} w(x,z).f(z)} \tag{1}$$

where $f(y)$ is the fitness of an agent y and $w(x,y)$ is the weight of the edge between x and y.

We also incorporate a second update mechanism. The motivation for its inclusion is as follows. Following several iterations of learning from local neighbours, each community is likely to be in a state of equilibrium—either total cooperation or total defection. Agents in these groups are receiving the same reward as their immediate neighbours. However, neighbouring communities may be receiving different payoffs. An agent that is equally fit as its immediate neighbours may look further afield to identify more successful strategies.

$$\frac{w(x,y).f(y)}{\Sigma_{z \in adj(adj(x))} w(x,z).f(z)} \tag{2}$$

where again $f(y)$ is the fitness of agent y and now (w,z) refers to the weight of the path between w and z. We use the product of the edge weights as the path weight. Note that in the second rule, we don't choose the agents in proportion to their edge weight values; we instead consider the complete set of potential in the extended neighbourhood. In this way all agents in a community can be influenced by a neighbouring cooperative community.

3.4 Noise and Dramatic Change

The first type of environment change we explore is the introduction of noise to the environment. Following every generation, the agent strategies are subjected to noise. For some agents, the first parameter of their strategy is changed from cooperation to defection and vice-versa. The second parameter in their strategy, the threshold (τ) is also subject to change; we merely reassign the value to be $1 - \tau$.

The second type of environmental change we model is that of a dramatic environmental change. We achieve this by reversing the effects of the agents' behaviours. Prior to the dramatic change, a certain action (cooperation) is individually dominated but collectively optimal; the other action (defection) is individually dominant but collectively results in a poor outcome. Subsequent to the change, defection is now the collectively optimal action and cooperation is the individually selfish move.

The experiments we undertake are aimed to explore the following issues: the influence of community structure, the emergence of cooperation among generalised *tft* strategies, the influence of game length on such emergence, the robustness of cooperation to noise and the robustness to dramatic environmental change.

3.5 Parameter Settings

In our experiments we use a population of 800 agents; we allow simulations to run for 750 generations. Following each generation, the first learning rule is applied. Following every four generations (sufficient for community to reach an equilibrium), the second learning rule is applied.

We vary the noise level using the following noise levels: 1%, 5%, 10%, 15% and 25%. Dramatic change is modelled by reversing the game pay-offs. We vary the following parameters—the strength of the inter-community links and the length of the game.

In this work, the intra-community links are held constant with a value of one. The value of the inter-community links have a large influence. The lower the value, the more robust to defection the cooperative clusters should be. The length of the game also has a large influence on the outcome. As the game length increases, the threshold parameter becomes more important than the initial move.

4. EXPERIMENT RESULTS

4.1 Experiment 1: varying community structure and game length

In these experiments, we let the game length range from two to four. We alter the level of community structure by holding the intra-community links at one and varying the inter-community links. We set the inter-community links to be 0.1, 0.3, 0.8 and 1.0.

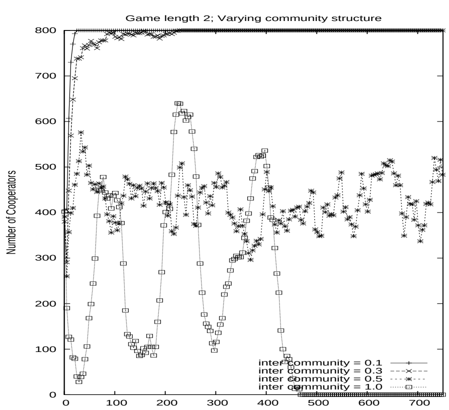

Figure 2: Game Length = 2; Varying community structure

Figures 2,3 and 4 show similar experiments for increasing game lengths. For each of the game lengths (2, 3 and 4) we explore the effect of the community structure. In Figure 2, we see that for high levels of community structure, cooperation emerges; for lower values, defection spreads. For the graph with no community structure (all edge weights equal to one), cooperation initially spreads but then quickly collapses. As the game length increases, the effect of the game length is more pronounced. In the final graph, with game length four, cooperation emerges in all cases. This is due to the effect of the *tft*-like strategies and the increased probability of reaching cooperative equilibrium states.

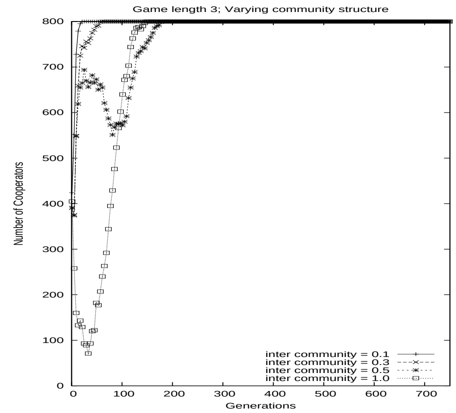

Figure 3: Game Length = 3; Varying community structure

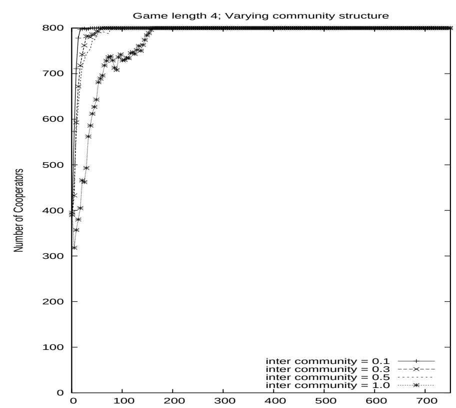

Figure 4: Game Length = 4; Varying community structure

4.2 Experiment 2: Introducing noise

In order to explore the effect of noise, we hold the length of game and the level of community structure fixed. We set the game length to be three and the inter-community links to be 0.1 reflecting a population with a high community structure. We see from Figure 5 that the population remains relatively robust to low levels of noise; up to 10 per cent noise can be tolerated with the population still remaining a high level of cooperation. Due to the presence of noise, defection will spread within a cluster. However, given the second learning rule, cooperation will quickly regain its position as the dominant strategy. This process of introducing noise causing an increasing in defection followed by the subsequent recovery of cooperation repeats itself. As the noise levels increase to higher values (15% and 25%), the population still does not tend to defection. The introduction of noises causes an immediate spread of defection but the cooperative clusters are relatively robust in the face of this invasion.

4.3 Experiment 3: Introducing dramatic change

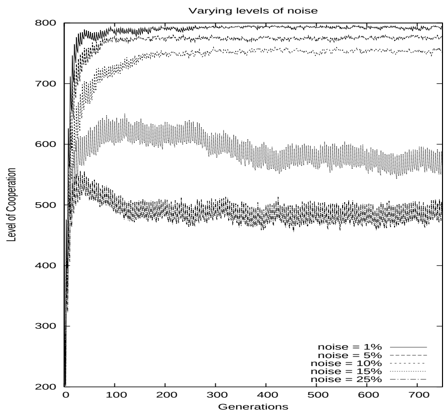

Figure 5: Varying levels of noise

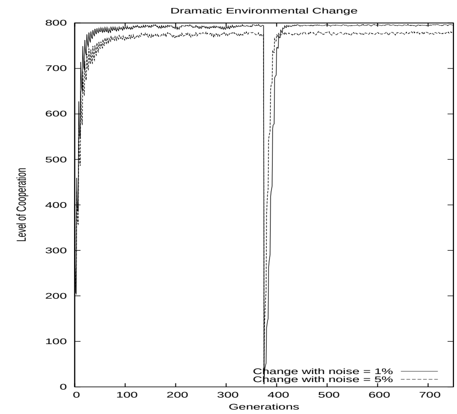

Figure 6: Changing Environment

The final experiment examines the effect of dramatically changing the agent environment. We reverse the payoffs for the game. We wish to explore whether or not our spatially organised population with a topology representing a high community structure can track the new socially benefical behaviour. Ideally the population should converge to cooperation and following the change in environment the population should converge to 'defection' in the new game. In reality this new 'defection' is really a form of cooperation (choosing the socially beneficial behaviour) and is counted as such in the graphs.

In Figure 6, we use a game length of three and a high community structure (inter community links fixed at 0.1). We run simulations for two levels of noise (1% and 5%). The dramatic environmental change occurs at generation 375. In the graph, we plot the number of agents adopting the socially beneficial behaviour (cooperation before generation 375, defection after generation 375). We see that the population is able to track the change in environment and quickly realise very high levels of socially beneficial.

Following the dramatic change, the fitness of the society drops dramatically as most agents are performing the individually rational yet collectively sub-optimal behaviour. Immediately prior to the change, a number of agents are 'defectors' due to the presence of noise; this spreads to their immediate neighbours so we have a few isolated clusters of defectors. Following the change these clusters are now performing optimally and this behaviour spreads quickly through the population due to the second learning rule.

5. CONCLUSION

In these experiments, we show that with a sufficiently high level of community structure, cooperation can emerge in a population of generalised tit-for-tat agents playing in a N-player social dilemma. Moreover, as the length of the game increases, there is less need for the community structure to encourage cooperation.

We also show that the spread of cooperation can be robust to the introduction of noise. Higher levels of noise do not necessarily drive the population to a state of total defection.

Finally, given the presence of a small degree of noise, high levels of socially beneficial behaviour can be acheived even with the introduction of extreme environmental change.

6. REFERENCES

[1] R. Axelrod. *Evolution of Cooperation*. Basic Books, 1985.

[2] R. Boyd and P. Richerson. The Evolution of Reciprocity in Sizable Groups. *Journal of Theoretical Biology*, 132:337–356, 1988.

[3] D.Curran, C.O'Riordan, and H.Sorensen. Evolutionary and lifetime learning in varying nk fitness landscapes with changing environments. In *AAAI*, 2007.

[4] S. Lozano, A. Arenas, and A. Sanchez. Mesoscopic structure conditions the emergence of cooperation on social networks. *arXiv:physics/0612124v1*, 2006.

[5] J. A. R. Marshall and J. E. Rowe. Viscous populations and their support for reciprocal cooperation. *Artificial Life*, 9(3):327–334, 2003.

[6] M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Physics Review E*, 69, 2004.

[7] M. Nowak, R. May, and S.Bonhoffer. More Spatial Games. *International Journal of Bifurcation and Chaos*, 4(1):33–56, 1994.

[8] C. O. Riordan, J. Griffith, J. Newell, and H. Sorensen. Co-evolution of strategies for an N-player dilemma. *Proceedings of the Congress on Evolutionary Computation*, June 19-23 2004.

[9] F. Santos and J. Pacheo. Scale-free Networks Provide a Unifying Framework for the Emergence of Cooperation. *Physical Review Letters*, 95(9), August 2005.

[10] J. Wu and R. Axelrod. How to Cope with Noise in the Iterated Prisoner's Dilemma. *Journal of Conflict Resolution*, 39(1):183–189, March 1995.

[11] Z. Wu, X. Xu, Y. Chen, and Y. Wang. Spatial prisoner's dilemma game with volunteering in newman-watts small-world networks. *Physical Review E*, 71:37103, 2005.

[12] X. Yao and P. J. Darwen. An experimental study of N-person iterated prisoner's dilemma games. *Informatica*, 18:435–450, 1994.