Evolution of Coordinated Behaviour in Artificial Life Simulations

Maud Gibbons, Colm O'Riordan, Josephine Griffith
Information Technology, National University of Ireland, Galway
{m.gibbons11, colm.oriordan, josephine.griffith}@nuigalway.ie

ABSTRACT

This paper describes research and results in evolutionary game theory. In particular we explore the evolution of mobility with coordination among agents inhabiting a virtual world where interactions are modelled using the Prisoner's Dilemma. We adopt an evolutionary game theory approach and consider a spatially structured population where a lattice grid is adopted. Agents' actions and movements in the game are controlled by their genotype which is subject to evolutionary pressure; fitness is calculated as the payoff that is accumulated by the agents in the simulation. A number of experiments are presented which show that agents evolved a form of coordination during the evolutionary simulation. Experiments detailing the effect of varying the game matrix and changing the initial clustering of the agents are also discussed.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Multiagent systems*; J.4 [**Computer Applications**]: Social and Behavioural Sciences

General Terms

Theory, Experimentation

Keywords

Coordination, artificial life, prisoner's dilemma, evolution

1. INTRODUCTION

In a range of domains in modern computing (multi-agent systems, distributed systems, peer to peer networks etc.), issues arise regarding coordination of nodes/agents in the system. Oftentimes, it is desirable to allow autonomy of agents but this renders centralised control and potential optimisation difficult. In order to fully understand these systems, there is a need to explore emergent coordination in systems. Many bodies of work have focused on cooperation and its emergence. Nowak and May [1] first presented the idea that cooperation can be maintained if cooperative strategies play with each other on a spatially structured model where agents played with their neighbours. This domain has since been well studied and many more ideas, theories, systems and experimental results have been presented. However, the area of coordination has been much less studied and there are relatively few studies that address spatial aspects of coordination among agents located in a virtual world.

Begun by John Maynard Smith, evolutionary game theory has been studied since the 1980s where ideas from evolutionary theory have been adopted into game theory [2]. We adopt the most oft-studied game in evolutionary game theory the classical Prisoner's Dilemma [3] as our interaction model. The game has attained popularity as it captures nicely the conflict between individually rational choices and those made for the common good. There have been many works exploring aspects of the prisoner's dilemma including work on evolving strategies [4], exploring the effect of spatial constraints including grids [1] and more complex topologies [5]. However, in many social scenarios that we may wish to model, participants often have more control over their interactions.

More recent studies have considered the movement of individuals as an effective means of promoting cooperation. Helbing and Yu [6] present a 'migration' model of a set of spatially organised mobile agents playing the Prisoner's dilemma game. Agents can consider moving to a different location in a specific 'migration range' based on whether the new location is in a better 'neighbourhood'. Ichinose et al [7] investigate the coevolution of migration and cooperation. Agents play an N-player Prisoner's Dilemma game following which they will move to a location in their (Von Neumann) neighbourhood based on a unique probability vector. Agents, both cooperative and non-cooperative, are evolved to collectively follow or chase the cooperaters. The authors highlight the importance of flexibility in the direction of migration for the evolution of cooperation.

This work differs from much previous research in that agents are afforded a more expressive form of dynamism, in that, agents can more explicitly chase desirable opponents and can also avoid adverse interactions. We are less focused on the evolution of cooperation in these games that in the emergence of coordination among the agents which can facilitate cooperation. Our model is more inspired by work in artificial life where moving populations of agents have been used to explore a range of phenomena. This paper describes a number of experiments using evolutionary computation to evolve behaviours in populations of agents interacting in a

ACM womENcourage'15, 24 - 26 September, Uppsala, Sweden.

	C	D
C	(3,3)	(0,5)
D	(5,0)	(1,1)

Table 1: Payoffs for Prisoner's Dilemma

simulated world. We explore the outcomes in the population and demonstrate the emergence of cooperation in some scenarios and coordination across a range of outcomes. The main contributions are as follows: we present a model for evolving coordination in a simulated world of agents playing the prisoner's dilemma; we present results showing that coordinated movement is indeed evolved; we also show the effect of clustering in the emergence of cooperation.

2. EXPERIMENTAL SET UP

This section presents an overview of the simulation set up. We describe the simulated world, the agents and their behaviours, and the evolutionary set up. The agents inhabit a simple grid-like world where each cell on the grid can host one agent. Each agent may interact with agents located in a neighbouring cell where the notion of a Moore neighbourhood is used. Each agent may also view agents located in this neighbourhood and can move to an adjacent cell.

Each agent is represented by a genotype comprising of seven bits. The first bit determines how the agent will interact in the game: cooperate or defect. The remaining bits determine how an agent will move. If an agent encounters a cooperator, they have a set of potential actions. These actions are as follows: remain where they are, move randomly, follow the cooperator or flee from it. Similarly these potential actions are mirrored for when an agent encounters a defector. The final two bits are used to determine actions for when an agent encounters both a defector and a cooperator. The actions are: flee from both; follow both; follow the cooperator and flee from the defector and the converse action (flee from the cooperator and follow the defector). During simulation, each potential action of an agent is determined by the genotype.

Each interaction involves pairs of agents participating in the Prisoner's Dilemma, which is a two player zero sum game where each player has a choice: to either cooperate or defect. If both agents cooperate, then both agents receive the payoff for mutual cooperation. However, there is a temptation to defect against a cooperator and receive a high payoff (temptation to defect). However, if both defect, they both receive a payoff worse than that received for mutual cooperation. Because betraying a partner offers a greater reward than cooperating with them, all purely rational, self-interested prisoners would betray the other, and so the only possible outcome for two purely rational prisoners is for them to betray each other [3]. Table 1 shows one commonly adopted game matrix where the values pairs refer to the payoffs received by the row and column player respectively.

The strategies that the population of agents can adopt were explored by Genetic Algorithms (GA). GAs offer an approach to search a complex space, and update agent behaviour based on a fitness score. Each agent is represented by an array of seven values, their genotype; this code determines the actions of the player; their phenotype. Each generation involves allowing players to move according to their genotype, interact with any neighbours following movement. Each interaction causes a change to their score. After a number of iterations, the accumulated score is taken as a measure of fitness. Agents for the next generation are initially selected based on this fitness score (the higher the better). Agents are also subject to GA operators of crossover and mutation (at pre-defined rates) to allow for more diversity in the population. The agents are evolved in this manner over a number of generations, and we measure levels of cooperation and the evolved behaviours over the evolutionary run.

There are a number of research questions under examination in this work. We are particularly interested in exploring whether coordinated movement can be evolved where agents evolve to cluster together to promote cooperation or not. We explore a number of research questions:

- Given the ability to evolve movement together with action in the game, will cooperation emerge?
- Given the ability to evolve both movement and actions, will coordination emerge?
- How robust will be any evolved cooperation to changes in the payoff matrix?
- If agents are clustered together (to different levels) at the start of generations will cooperation be favoured?

3. EXPERIMENTAL RESULTS

In these experiments, the following parameters are used in each experiment. For these experiments, a 50 x 40 grid world is created and populated with 100 prisoners, initially the cooperator to defector ratio is 1:1. Simulations are run for 50 generations, and prisoners will take 200 turns each generation. The simulator will be run at least 10 times for each experiment. The payoff matrix used is that described earlier, unless otherwise noted. Tournament selection of size two is adopted, crossover of 70% and 1% mutation. Any other parameters specific to a particular experiment is documented for that experiment.

3.1 Experiment 1: Evolution of Coordination

The aim of this experiment is to explore and analyse behaviours in the population of agents. We wish to observe the evolution of the prisoners' behaviours (both for the interaction and their movement) with evolutionary pressure being placed on both behaviours. In this experiment, we record the fitness of the population over time and the behaviours of each prisoner at each generation.

We include a number of plots to illustrate the evolution of behaviours reflecting coordination among agents. Similar coordinated behaviour was found for both cooperators and defectors. The plots show which behaviours evolved given some of the potential scenarios a cooperator may find themselves in. Figure 1 shows the evolved behaviours when a cooperator encounters a cooperator. The behaviour 'follow' dominates quite quickly, showing that for cooperators, they learn to follow each other and coordinate their movement in order to gain high rewards for increased interactions involving mutual cooperation.

In Figure 2, we again see that cooperators quickly coordinate their movements for scenarios where a defector is encountered. The 'flee' behaviour quickly dominates as agents

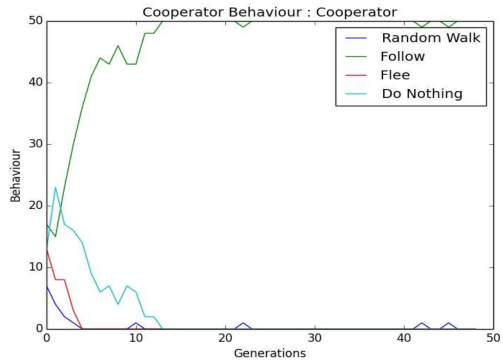

Figure 1: Evolved behaviours for cooperators when encountering a cooperator

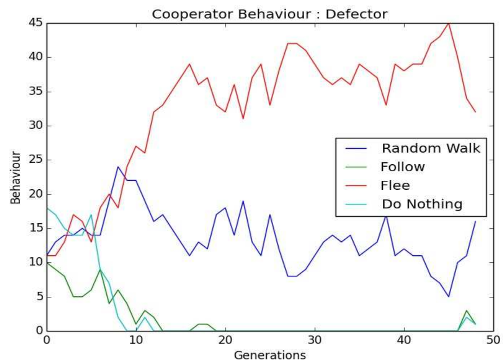

Figure 2: Evolved behaviours for cooperators when encountering a defector

attempt to avoid exploitation by those who defect. The plot also shows the movement 'random walk' where the agents chose a random move upon seeing a defector does not die out. There are two reasons for this; firstly, the behaviour is reasonably good for a cooperator encountering a defector as it will in most cases lead to the agent fleeing from the defector and secondly following a number of generations for this run, the defectors have died out so there is no longer a selective pressure on these movements.

When a cooperator encounters both cooperators and defectors, the results show that they quickly evolve to follow cooperators and flee from defectors. This leads to agents that cooperate when clustering together resulting in a form of emergent coordination. These graphs show the behaviours of cooperators for just one run of the simulation where cooperation prevailed.

As agents are placed randomly in the grid at the start of each generation, there can be an impact on the outcome of the evolutionary run. Different initial positioning of agents can lead to a benefit for either cooperators or defectors and can affect the evolutionary trajectory. In order to get a more complete view of the evolution, we run the simulation a number of times and present results on the outcomes. For 20 simulations, the evolutionary process resulted in cooperative outcomes 13 times (65%) and 7 times (35%) in non-cooperative outcomes. Notwithstanding, the outcome of the overall simulations, the same behaviours were evolved across all scenarios, agents evolved to move towards cooperators and to flee from defectors. This was true for both cooperators and defectors as both gained in fitness from interacting with cooperators and both achieved reduced fitness from defector interactions.

3.2 Experiment 2: Varying the Payoffs

The payoff matrix used in the prior experiments have been used in many previous works. However, we can change the payoffs while still maintaining the dilemma. For the prisoner's dilemma to hold, the temptation to defect (T) must be greater than the reward to mutual cooperation (C), which in turn must be greater than the reward for mutual defection (D) which must be greater than the sucker's payoff (S) when one cooperates and the opponent defects. We can vary this value while maintaining the above constraints. In this experiment, we vary the score available for mutual defection. By reducing the score, there is more advantages to be gained through cooperation and conversely if we increase the score for mutual defection to be close to that for mutual cooperation, defection should be favoured. We re-run the same experiments in section 3.1, but with different scores for mutual defection.

We tabulate the number of outcomes that result in cooperation as we change the score for mutual defection, running each simulation 20 times. This is summarised in Table 2. As can be seen from the table, decreasing the payoff to

Mutual Defection Score	Cooperative Outcomes
0.5	18
1.0	13
1.125	13
1.1815	10
1.25	9
1.5	7
2.0	6
2.5	6
2.99	6

Table 2: Number of cooperative outcomes for different mutual defection rewards

0.5, increases the number of cooperative outcomes and as the payoff for mutual defection tends to 3, (the payoff for mutual cooperation), the number of cooperative outcomes decreases. In these simulations, the same evolved coordinated movement was witnessed for lower scores for mutual defection; agents learned to follow cooperators in increase their fitness and to flee defectors. As the payoff for mutual defection increased, defectors gained little from following cooperators over defections or from fleeing from defectors, and different behaviours were evolved. This helps explains why there were still several cooperative outcomes. Cooperators evolved to follow cooperators and avoid defectors. Defectors did not have the same selective pressure on these behaviours and were content to have interaction with fellow defectors.

3.3 Experiment 3: Clustering Agents

The aim of this final set of experiments is to explore and analyse the effects of clustering on the outcome of the simulation. We hypothesise that it is the clustering of cooperators that leads to cooperative outcomes. In the previous experiments, cooperators learn to follow other each other leading to a clustering of these cooperators. In order to more clearly explore the effect of this emergent coordination and clustering, we enforce a level of clustering at the start of each generation so as to control the amount of cooperator interactions.

In these experiments, we place agents with a level of randomness but control the mix between cooperators and defectors. In an extreme case, we place all the cooperators and no defectors in a certain region of the grid (top left quarter of cells) and place the defectors in the remainder of the grid in a random manner. This ensures that cooperators will be clustered together thereby increasing the number of mutually cooperative interaction and reducing the probability of exploitation by defectors. We then vary the level of clustering by swapping some cooperators from the cooperative cluster with the surrounding defectors. We run a set of experiments for each of these initializations. We return to using the standard payoff matrix as defined in Table 2. In this experiment the ratio of cooperative to non-cooperative simulation outcomes will be measured after each change in the level of clustering. Having run the simulator multiple times (50) we see that cooperation, with cooperative prisoners clustered together in the top left of the grid, is highly robust.

Defectors Swapped	Cooperative Outcomes
0	98%
5	90%
10	86%
15	60%
20	50%
25	18%

Table 3: Cooperator success with variation in the levels of Mixing

Table 3 shows us the proportion of simulations that resulted in a population of defectors as we changed the level of mixing between the cooperators and defectors. In a homogeneous environment, a simulation will result in a cooperative victory 98% of the time, with defection only winning out in 2% of the simulations; this is a dramatic increase compared to previous experiments. The cluster of cooperators itself is also fairly robust; it can easily withstand the defectors introduced at mixing levels of 5 and 10 with the vast majority of simulations resulting in a cooperator win. When the number of defectors injected into the cluster is increased we can increase the probability that a simulation will result in a defector win.

4. CONCLUSION

This paper presented a series of results in artificial life simulations of cooperation and coordination. We show that by allowing agents to evolve both their actions and their movements, we can induce both cooperation and coordination in an artificial world. The results show that agents learn to choose to follow cooperators and flee defectors; this leads to an emergent form of coordination for cooperators who attempt to follow each other while fleeing defectors which increases their frequency of cooperative interactions. The latter experiments show the effect of changing the payoff matrix to reflect different balances between mutual cooperation and mutual defection. The final experiment shows explicitly the effect of the coordination of cooperative agents, which is emergent in the early experiments. In the final experiment, in an effort to illustrate its effect we explicitly enforce this clustering and show it has a huge impact on the outcome.

In conclusion, this paper illustrates our success in the evolution of mobility with coordinated behaviour. A population of cooperative agents has been evolved with the ability to coordinate themselves spatially, forming clusters, giving them a competitive edge over the self-interested, non-cooperative defectors.

5. CURRENT & FUTURE WORK

Extensions to this work has included making changes to the simulator to move from a simple grid to a toroidal shaped world, where the sides and edges of the grid connect. This eliminates any potential edge effects that may restrict the movement or influence the evolutionary trajectory of the agents. An algorithm was also developed to more formally analyse the levels of coordination present in the population and investigate the emergent hypothesis that cooperator clustering leads to a cooperative outcome. It explicitly measures the frequency, size, and cohesion of clusters in the grid at various time steps during a simulation.

Future work will involve attempting to predict the outcome of a simulation based on examination of the initial positioning of agents. It will also involve exploring other interaction models to capture more complex interactions. In particular we will explore an extension to the Prisoner's Dilemma which allows participants to abstain from playing the game.

6. ACKNOWLEDGEMENT

We would like to acknowledge the support of NUI Galway's Hardiman Research Scholarship.

7. REFERENCES

[1] Martin A. Nowak and Robert M. May. Evolutionary games and spatial chaos. *Nature*, 359(6398):826–829, 1992.

[2] John Maynard Smith. *Evolution and the theory of games*. Cambridge University Press, Cambridge; New York, 1982.

[3] Robert M. Axelrod. *The evolution of cooperation*. Basic Books, 1984.

[4] Paul J. Darwen and Xin Yao. On evolving robust strategies for iterated prisoner's dilemma. In *Progress in Evolutionary Computation*, pages 276–292. Springer Berlin Heidelberg, 1995.

[5] Menglin Li and Colm O'Riordan. The effect of clustering coefficient and node degree on the robustness of cooperation. In *2013 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2013.

[6] Dirk Helbing and Wenjian Yu. The outbreak of cooperation among success-driven individuals under noisy conditions. *Proceedings of the National Academy of Sciences*, 2009.

[7] Genki Ichinose, Masaya Saito, and Shinsuke Suzuki. Collective chasing behavior between cooperators and defectors in the spatial prisoner's dilemma. *PLoS ONE*, 8(7):e67702, 07 2013.