Cooperation Without Memory or Space: Tags, Groups and the Prisoner's Dilemma

David Hales

Department of Computer Science, University of Essex, Colchester, Essex, UK
daphal@essex.ac.uk

Abstract. A Recent [14] model demonstrated that image scoring produces high cooperation between strangers in the Prisoner's Dilemma (PD). Here we outline a simpler approach in which players – which are either pure cooperators or defectors – can sustain cooperation with strangers by biasing game interactions towards others with the same tags (arbitrary bit strings representing "cultural markers" [11]). In our model there is no requirement for knowledge of past performance or recognition of individual players. Unlike spatial models [13] reproduction of strategies is population wide. Contrary to previous tag models [15] cooperation is demonstrated in the single round game.

1 Introduction

In modern and complex social worlds, individuals are required to interact cooperatively with many strangers using limited knowledge and bounded rationality. But why do strangers cooperate? Here we discount those situations in which cooperation is possible without trust and examine that subset of cooperative interactions that follow the form of a social dilemma [10]. We formalise this kind of encounter using the ubiquitous form of the two player single round Prisoner's Dilemma game. We demonstrate empirically, for the first time, that the biasing of game interaction towards agents sharing identical tags (arbitrary markings represented as bit strings) is sufficient to produce high levels of cooperation in the single round PD when agents are boundedly rational optimisers. Interestingly, this process can be visualised as the formation and dissolution of "groups" that share the same tags in a non-physically extended abstract "tag space". Firstly we introduce the Prisoner's Dilemma game, then we outline some existing theories which attempt to explain the emergence of cooperation within such a game between evolutionary optimisers. We then introduce our minimal tag based model (the TagWorld) and give the results obtained. Finally we discuss the significance of the results and their possible applicability to human societies.

2 The Prisoner's Dilemma

The Prisoner's Dilemma (PD) game models a common social dilemma in which two players interact by selecting one of two choices: Either to "cooperate" (C) or "defect" (D). From the four possible outcomes of the game payoffs are distributed to the individuals. A reward payoff (R) and a punishment payoff (P) are given for mutual cooperation and mutual defection respectively. However, when individuals select different moves, differential payoffs of temptation (T) and sucker (S) are awarded to the defector and the cooperator respectively. Assuming that neither player can know in advance which move the other will make and wishes the maximise her own payoff, the dilemma is evident in the ranking of payoffs: $T > R > P > S$ and the constraint that $2R > T + S$.

Although both players would prefer T, only one can attain it. No player wants S. No matter what the other player does, by selecting a D move a player ensures she gets either a better or equal payoff to her partner. In this sense a D move can't be bettered since playing D ensures that the defector can not be suckered. This is the so-called "Nash" [8] equilibrium for the single round game. It is also an evolutionary stable strategy [12] for a population of randomly paired individuals playing the game where reproduction fitness is based on payoff. But the dilemma remains, if both individuals selected a cooperative move they would both be better off. But many societies (human and animal) appear to have solved (at least some) dilemmas similar to the PD. How can this be explained by purely evolutionary mechanisms?

3 Evolutionary Extensions

Evolutionary selection favours selfish individual replicators. When collections of these replicators form groups it's possible for them to co-ordinate their behaviour in ways which would make global optimisation possible. The kinds of behaviours that make this possible include, cooperation, altruism and specialisation. All of these are observed in animal and human societies. But evolutionary selection does not seem to offer an explanation for these behaviours. To address this problem three extensions of natural selection have been proposed: kin selection [9], group selection [17] and reciprocal cooperation [2]. Although each offers explanations of some of the kinds of the social behaviours of interest neither seems to offer a general framework applicable to human or artificial social systems. Kin selection only applies to highly genetically related individuals, group selection in its simplest form is fundamentally flawed (selfish individuals within the group benefit relative to altruists) and reciprocal cooperation does not explain true altruism (i.e. cooperative behaviour in the single round PD). Neither does it scale-up well to large groups due to the cognitive demands from the requirement that all interactions be on-going with recognisable individuals and associated memory of past interactions.

4 The TagWorld

Agents are represented as fixed length bit strings (of length L+1) comprising a tag of length L bits and a single strategy bit. The strategy bit represents a pure strategy, either unconditional cooperation or unconditional defection. Initially the population of agents are set to random bit strings (with each bit decided by a fair coin toss). The following evolutionary algorithm is then applied:

```
LOOP some number of generations
  LOOP for each agent (a) in the population
    Select a game partner agent (b)
    Agent (a) and (b) invoke their strategies receiving
      the appropriate payoff
  END LOOP
  Reproduce agents in proportion to their average payoff
END LOOP.
```

In each generation each agent (a) is selected from the population (of size N) in turn. A game partner is then selected. Partner selection entails the random selection of another agent (b) from the population such that (a) ≠ (b) followed by a comparison of tags with agent (a). If the tags are identical a game interaction takes place otherwise (b) is returned to the population without game interaction. If (b) was returned to the population without interaction a second (b) is selected at random from the population and its tag compared with (a). This process is repeated until an agent (b) is found which has an identical tag to (a) or an upper limit F of selections has been reached. If this upper limit is reached then game interaction is forced on the next randomly chosen agent. Consequently (a) will always find a partner even if its tag does not match any other agent because an agent which can not find a matching partner will eventually exhaust its upper limit F of refusals and then be forced to interact with a randomly chosen partner.

During game interaction (a) and (b) invoke their strategies and receive the appropriate payoff. After all agents have been selected in turn and played a game a new population is asexually reproduced. Reproductive success is proportional to average payoff. The entire population of N agents is replaced using a "roulette wheel" selection method [5]. Equation 1 and inequality 2 outline this method. Equation 1 gives the total average payoff for the entire population where a_i is the *ith* agent from the population, *ap(a)* is the average payoff obtained by agent *a*, and N is the size of the population. The inequality 2 specifies an agent a_x to select for reproduction from the population where *x* is the smallest integer that satisfies the inequality and *rnd(0..tap)* is a uniformly randomly selected value in the range *0..tap*. The inequality is satisfied N times with a different random value. Each time, *x* gives the index of an agent to reproduce. Using this method the probability that an individual will be reproduced into the next generation is proportional to average payoff.

$$tap = \sum_{i=1}^{N} ap(a_i) \tag{1}$$

$$rnd(0..tap) \leq \sum_{i=1}^{x} ap(a_i) \tag{2}$$

Mutation is applied to each reproduced player with probability M = 0.001. This low value indicates the assumption that agents rarely change their strategy. Since there are 100 agents in the population, we would expect one strategy bit to change over 10 generations (on average). Mutation takes the form of flipping each bit of each player with probability M. Consequently *tags and strategies* are mutated in reproduced agents.

The PD payoffs are parameterised over T (the temptation payoff for defectors over cooperators) such that T > 1. The reward R for mutual cooperation is 1. The punishment P for mutual defection and the sucker payoff S for cooperation with a defector are both 0.0001. This value was selected because it was small but greater than zero (indicating a very small chance for agents, with Sucker or Punishment payoffs, of reproduction). If a small value is added to P (enforcing T > R > P > S) results are not significantly changed.

5 Results – High Cooperation

A set of runs to 100,000 generations with a population of size N = 100 agents was executed for various values of T and L. The maximum number of refusals was fixed at F = 1000. This high value of F means that it is unlikely that agents will not be paired with other matching agents (if they exist) in the population. For the purposes of analysis cooperation was characterised as the proportion of mutually cooperative interactions occurring over all generations. This figure was calculated by counting the number of games in which both agents cooperated for the whole run (of 100,000 generations) and then dividing by the total number of games played. Thus the level of cooperation for a single run is derived from the results of 10^7 individual games.

Figure 1 shows results for various values of L and T graphically. Each bar represents an average of 5 independent runs. As the values of L are increased, and T are decreased, cooperation increases monotonically. As can be seen, where L >= 32 very high levels of cooperation are obtained for all values of T.

The results obtained indicate that very high levels of cooperation can be sustained between optimising agents in the single round PD via simple tag biasing. There is no requirement for knowledge of past performance or recognition of individual agents (i.e. the other agents may be viewed as strangers). Unlike some spatial models which have demonstrated cooperation in the single round game [13] reproduction is population wide. The next section explains this high cooperation in terms of group formation.

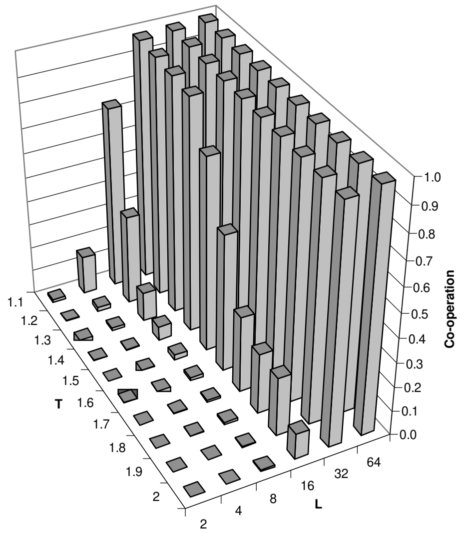

Fig. 1. Cooperation for various values of L (number of tag bits) and T (the temptation payoff) are shown. Cooperation is measured as the proportion of mutually cooperative games over 100,000 generations. Each point is an average of 5 runs. The entire chart represents 3 x 10^9 individual game interactions.

6 Group Formation and Dissolution

The tag space can be visualised as an L-dimensional hyper-cube with corners representing unique tag values. Agents sharing a tag, share a corner. Mutation produces movement between corners. Game interaction is therefore taking place in an abstract "tag space". Cooperative groups sharing matching tags will form in corners of the hyper-cube. These groups will outperform non-cooperative groups and hence tend to increase in size over generations. However, if mutation introduces defecting agents into a cooperative group they will tend to outperform the cooperators within the group (by suckering them). From this the seeds of the destruction of the group are planted, since as the number of defectors increases within a group the overall fitness of agents within the group decreases. Other more cooperative groups (if they exist) will tend to expand. While this process is occurring, mutation of tag bits will produce a slow migration of agents between corners of the hyper-cube, possibly founding new groups in previously empty corners.

Figure 2 is a visualisation of the process over time taken from a single run. Each line on the vertical axis represents a corner of the hyper-cube (i.e. unique tag value). The horizontal axis represents time in generations. If no agents have a particular tag value in a given generation then the line is left blank (white). Alternatively, if a corner contains all cooperative agents then the line is light grey. For a mixed group in which there are both cooperators and defectors the line is dark grey. For an entirely defective group the line is black. Examination of figure 2 shows the time evolution of groups in tag space. Initially cooperative groups (light grey lines) become invaded by defectors producing mixed groups (darker grey) which very swiftly become entirely defective (black) and then quickly go extinct (white).

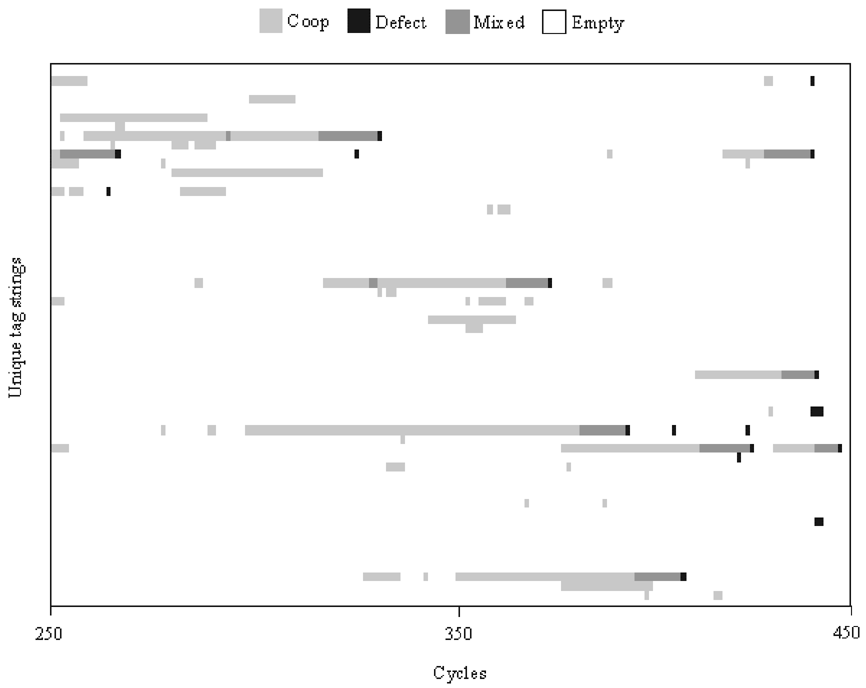

Fig. 2. Visualisation of 200 cycles (generations) from a single simulation run showing cooperative groups coming into and going out of existence. Each line on the vertical axis represents a unique tag value (of which only a subset is shown). If all agents sharing a tag value are cooperative then the line is light grey. If all agents are defectors then the line is coloured black. A mixed group is shown as dark grey. The horizontal axis represents time in cycles (generations). Here L = 8 and T = 1.1.

7 Cooperation from Complete Defection

Although high levels of cooperation are demonstrated over many contiguous generations, starting from the random initialisation of agents, these results do not indicate if a society can recover cooperation from a state of complete defection. In order to test this, experiments were conducted in which the initialisation of agents was modified so that all agent strategy bits were set to defection. Figure 3 shows a set of runs for various values of N (population size) against the number of generations before mutual cooperation emerges. These empirical results are compared to a simplified analytical (probability based) treatment given in equation 3.

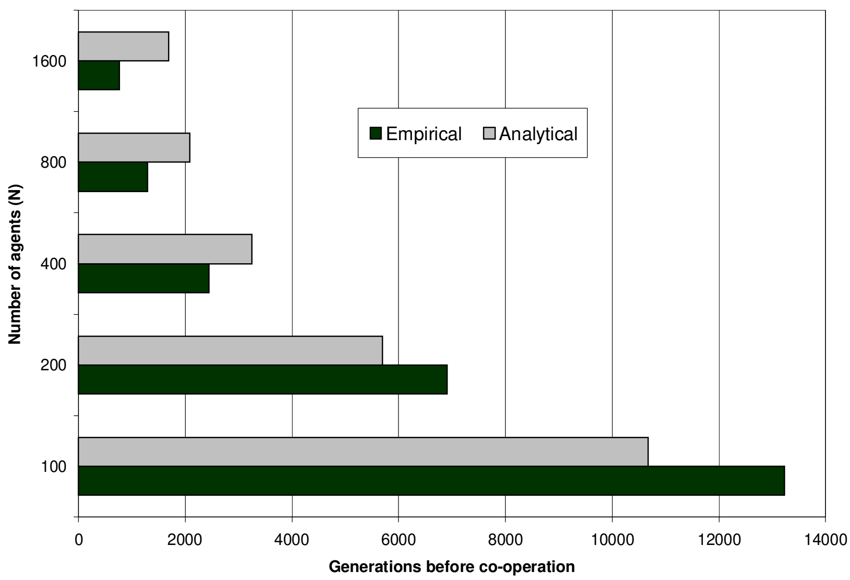

Fig. 3. Number of generations before mutual cooperation emerges. Here a comparison is made between the analytical model (given in equation 3) and the empirical results. Each bar for the empirical test comprises an average of 10 runs. Tag size L = 32 for all runs shown.

$$ang(n,m) = \frac{1}{\frac{2}{n-1}\left(1-(1-m)^{n} - nm(1-m)^{n-1}\right)} \qquad (3)$$

Equation 3 gives the expected average number of generations required (from an initial society of all defectors) before two (and only two) cooperative agents perform a game interaction. This is dependent on the population size N and the mutation rate M. In all of the 50 runs used to form the empirical results in figure 3 it was observed that high cooperation immediately (in the next generation) followed the first occurrence of a mutually cooperative encounter between two agents. It was also observed that drift over the tag bits tended to lead a society of all defectors toward sharing the same tag bits. These empirically observed (rather than derived) phenomena were used to simplify the analytical treatment: we assume all agents share the same tag bits and high cooperation starts when two agents mutate to cooperators and game interact.

In figure 3 the predictions of equation 3 are compared empirically with actual runs. As can be seen the analytical treatment tends to under-estimate the number of generations before cooperation emerges for low population sizes and over-estimate for large population sizes. It can be hypothesised that this is due to the simplification in equation 3. Drift will rarely produce populations in which all agents have identical tags. When the population is split between different tag values the simplification will under-estimate the number of generations required. This is because two cooperators may be produced but they may not share the same tag and hence will not game interact. Conversely when the population size is large it will often be the case that more than two agents will be mutated to cooperators. A more detailed analysis would need to capture the dynamics of the tag bits over the population and the effect of those dynamics on producing cooperative interactions.

8 Discussion

In a previous study of tag based partner selection in the PD, Riolo [15] concluded that tags produced little increase in cooperation in the single round game. In his work a tag was represented as a single real number attached to each agent. The abstract topology of the tag space was therefore one dimensional. The matching of tags was based on a probabilistic function applied to the distance between two tags, meaning that agents with similar but not identical tags could engage in game interactions. The number of refusals allowed before forced interaction was low (50 refusals in a population of 200 agents). Additionally a fitness cost was attached to each refusal made by an agent (although this was reduced to zero in an attempt to get high cooperation in the single round game). Under these conditions it was demonstrated that high cooperation emerged when agents engaged in the Iterated PD (IPD) but not in the single round PD game. The interpretation placed on this previous study was of agents representing animals, searching for game partners and evolving genetically. In the work presented here allowable refusals is high (10 times the population size) and there is no associated cost. Also the tag is represented as a bit string which must match *exactly* with a partner for game interaction to be selected by an agent. Under these assumptions a different kind of tag space topology is possible and high cooperation is produced in the single round game.

Intuitively it would seem that the *exact tag matching* constraint is not necessary to produce high cooperation in all cases. For example, consider a situation in which mutation was zero and two groups of agents existed in the tag space such that there was no inter-group game interaction. If one group consists of all cooperators and the other contains some defectors then the cooperative group would expand at the expense of the non-cooperative group. This would even hold if agents were applying some partial matching scheme - *so long as there is an interaction boundary between the two groups*. By "interaction boundary" is meant that some mechanism partitions the agents into strict game interaction groups. That games can only take place between individuals sharing a group.

The interpretation placed on this work is of agents representing a "bundle" or "complex" of culturally learned and transmissible behaviours, a so-called "meme-complex" [3], [7]. The assumption is that in a population of hosts meme-complexes which produced high utility (for their hosts) would be more likely to be copied (in proportion to their relative utilities). One key to such a process producing high altruism and co-operation is the packaging of the tag with the beneficent strategy or behaviour as a cultural unit. Allison in his theory of altruism [1] echoes this assumption with reference to the importance of "cultural packaging techniques". In the work presented here "packages" of tags and defective strategies do not dominate the population because such packages destroy the very groups that they are a part of.

In recent work Bowles & Gintis [4] give a detailed analytical treatment of the value of groups in the promotion of co-operation in the PD when binary social cues convey useful predictive information concerning a game partners strategy. However, they do not address the issue of *how* social cues come to be have such predictive utility. The results produced within the TagWorld society show that even simple mechanisms can produce this kind of correlation because groups which contain defectors quickly die out. In contrast to the examples used by Bowles & Gintis (who focus on racial groups) the mechanism which produces this quick extinction of non-cooperative groups within the TagWorld society requires that cultural interaction, in the form of individuals moving between groups easily, is high. It is this ability of individuals to quickly swap cultural groups, by taking on new tags from others, which drives the co-operation producing process. Strong group boundaries which prevent easy entry and exit from a group would hinder or even destroy co-operation forming by the process illustrated in the TagWorld.

The TagWorld was parameterised over a number of dimensions. A scan was made over a restricted part of the parameter space demonstrating that high co-operation is present over a that area of that space. However, other dimensions of the space have not been explored. The role of mutation and refusals would be an interesting area of investigation. Also, measures other than just co-operation, such as group sizes over time and migration rates between groups, would be of interest and could be used to elaborate an analytical model. This could link this work with patch based models of altruism developed within evolutionary biology [17]. Another interesting area of further investigation would be to make refusals an endogenous parameter encoded into each agent and able to evolve. In such a scenario would high refusal rates evolve? If so then the assumption of high refusals could be justified via endogenous evolutionary processes rather than as an exogenous assumption of the model.

Acknowledgements

This work would not have been possible without the help of, discussions with, and suggestions from, Prof. Jim Doran, Department of Computer Science, University of

Essex, Colchester, UK. I am also grateful to Rick Riolo (Program for the Study of Complex Systems, The University of Michigan, Ann Arbor, MI, USA) for his perceptive and helpful comments and suggestions on a earlier draft of this paper. This work was supported by an EPSRC research studentship award and is based on work for a forthcoming PhD thesis.

References

1. Allison, P.: The cultural evolution of beneficent norms. Social Forces, 71(2), (1992) 279-301.
2. Axelrod, R.: The Evolution of Cooperation. Basic Books, New York (1980)
3. Blackmore, S.: The Meme Machine. Oxford University Press, Oxford (1999)
4. Bowles, S. & Gintis, H.: Optimal Parochialism: The Dynamics of Trust and Exclusion in Networks. Santa Fe Institute Working Paper 00-03-017. Santa Fe, NM (2000)
5. Davis, L.:Handbook of Genetic Algorithms. Van Nostrand Reinhold, New York (1991)
6. Dawkins, R.: The Selfish Gene. Oxford University Press, Oxford (1976)
7. Dennett, D.: Darwin's Dangerous Idea. Simon & Schuster, New York (1995)
8. Gibbons, R.: A Primer In Game Theory, Harvester, New York (1992)
9. Hamilton, W.: Altrusim and Related Phenomena, Mainly in Social Insects. In: Annual Review of Ecology and Systemics 3 (1972) 193-232
10. Hardin, G.: The tradigy of the commons. In: Science. 162 (1968) 1243-1248
11. Holland, J.: The Effect of Lables (Tags) on Social Interactions. Santa Fe Institute Working Paper 93-10-064. Santa Fe, NM (1993)
12. Maynard Smith, J.: Evolution and the Theory of Games. Cambridge University Press, Cambridge (1982)
13. Nowak, M. & May, R.: Evolutionary Games and Spatial Chaos. Nature, 359, (1992) 532-554
14. Nowak, M. & Sigmund, K..: Evolution of indirect reciprocity by image scoring. Nature, 393, (1998) 573-557.
15. Riolo, R.: The Effects of Tag-Mediated Selection of Partners in Evolving Populations Playing the Iterated Prisoner's Dilemma. Santa Fe Institute Working Paper 97-02-016. Santa Fe, NM (1997)
16. Simon, H. A.: A Mechanism for Social Selection and Successful Altruism. Science, 250 (1990) 1665-1668
17. Wilson, E. O.: Sociobiology: The New Synthesis. Harvard University Press, Cambridge Mass (1975)