Selection of Cooperative Partners in n-Player Games

No Author Given

No Institute Given

Abstract. We address the problem of finding the appropriate agents to interact with in n-player games. In our model an agent only requires knowledge about the payoff and identification of its partners. This information is used to update a probability distribution over candidate partners. As such, our model is applicable in any situation, be it a cooperative dilemma or a game where a Nash Equilibrium is equal to a Pareto Optimal profile.

1 Introduction

Reputation management [5, 4], partner punishment [4], partner selection [2, 10], network structure [12, 8] are models put forward to explain or analyse the prevalence of cooperative agents in games where a dilemma is present. However, the extensibility of such approaches to any game is often not discussed, as the proposed solution only applies to a specific game [2, 9, 1].

As we use Game Theory to model the interactions between agents, our model makes direct use of the payoffs of a game in order to select the cooperative partners. The model consists of a probability vector maintained by each agent where each position represents the probability of selecting an agent as a partner to play a game. With this approach we are able to apply our model to any situation capable of being described as a game, with partner identification.

Since the agent has to find the best partner, the algorithm can be compared to a Cournot adjustment process [6] were players iteratively adjust their strategies to their partner responses. In this paper, an agent strategy remains constant but it adjusts its preferences towards more profitable or cooperative partners. Similar approaches to partner selection have been tackled in [2, 10] but they focused on a specific game such as Prisoner's Dilemma (PD). Here we study the problem of partner selection in n-player games. In this case, the assessment of responsibility for the outcome is more difficult to make, due to increased uncertainty of having $n-1$ partners instead of a single one.

2 Definitions

Game Theory is a tool to model interaction between agents. To this end, we consider that a population $\mathbb{P}$ of agents interacts accordingly to the rules of some

n-player game $\mathcal{G}$. The game describes the strategies available to players and the payoffs they obtain as a function of the strategies used. The game has a n-dimensions strategy space $\mathbf{S} = \mathbb{S}_1 \times \mathbb{S}_2 \times \ldots \times \mathbb{S}_n$ where agents can draw a strategy $s \in \mathbb{S}_i$ to play a game. The vector $\mathbf{s} = (s_1, \ldots s_n)$ represents a strategy profile of the n players involved in the game. The game also has n payoff functions, $u_i : \mathbf{S} \to \mathbb{R}$, with $i \in \{1, 2, \ldots, n\}$. The payoff functions are bounded and belong to $\mathbb{R}$. Let $\underline{u}$ be the lowest payoff and $\bar{u}$ be the highest payoff in game $\mathcal{G}$.

We aim at reaching a position where cooperative agents only interact between themselves. As cooperative agents we define those that form a strategy profile that maximises the average payoff of the players. We define the payoff obtained by such Pareto Optimal profile as follows:

$$u_P = \max_{\mathbf{s}} \sum_i \frac{u_i(\mathbf{s})}{n}.$$

For example, in a Public Good Game a cooperative agent is one that contributes to the common good, and in the Common Pool Resource game a cooperative agent does not over exploit the resource (see [7] for a specification of these games).

3 Model Description

A population $\mathbb{P}$ of agents is represented by a directed simple graph where a vertex represents an agent α and $w_{\alpha,\beta}$ is the label of an edge from α to β representing the probability of agent α interacting with β:

$$w_{\alpha,\beta} \geq 0,$$
$$\sum_{\beta} w_{\alpha,\beta} = 1.$$

3.1 Update policy

The edge weight update policy for agent α is a function defined as follows:

$$w_{\alpha,\beta}^{t+1} = \zeta(w_{\alpha,\beta}^t, u_\alpha)$$

where $w_{\alpha,\beta}^t$ is the edge weight before the game in which agent α participated, u_α is the payoff of the agent in the game. Index β varies through all neighbours of α.

The main focus of the work presented in this paper is the analysis of an update policy that meets the following two conditions:

Cooperative aggregation – Cooperative agents are mostly connected to each other. If α and the set of its cooperative neighbours B are part of a Pareto Optimal profile, then in the limit the sum of the probability of selecting only $\beta_C \in B$ should be 1:

$$\sum_{\beta_C \in B} \lim_{t \to \infty} w_{\alpha,\beta_C}^t = 1 \quad B = \{\beta : u_\alpha(\ldots, s_\beta, \ldots) = u_P\}. \tag{1}$$

Stability – The update policy must be robust in order to resist perturbations in the population and to be applicable to any n-player game. In the long run and in the absence of perturbations, weights must stabilise:

$$\lim_{t \to \infty} (w^{t+1}_{\alpha,\beta} - w^t_{\alpha,\beta}) = 0. \tag{2}$$

The edge weight update policy function is divided in two cases depending on whether an agent played the game with agent α or not.

Agent β played the game. A simple policy is to multiply the old weight by a factor that is inversely proportional to the distance between payoff u obtained by agent α and the Pareto Optimal payoff u_P if u is lower than u_P. If it is higher or equal, the edge weight remains the same. The rationale being there is no motive to decrease the probability of selecting the current partners. The definition is:

$$w^{t+1}_{\alpha,\beta} = \begin{cases} w^t_{\alpha,\beta} \dfrac{u - \underline{u}}{u_P - \underline{u}} & u < u_P \\ w^t_{\alpha,\beta} & u \geq u_P. \end{cases} \tag{3}$$

This rule by itself does not guarantee the condition in equation (1). Only combined with the rule for the case of agents that were not selected we achieve it. Regarding stability, this rule will keep weights unchanged if the payoff is greater or equal than u_P. Otherwise they will tend to zero as in the first case $w^t_{\alpha,\beta}$ is multiplied by a factor always less than 1. Either way, equation (2) is met.

Agent γ did not play the game. As we have just seen, the multiplicative factor used for agents that played the game implies that the weight of all agents that played will either stay the same or decrease. If they decrease, the difference must be distributed among the other edge weights. A simple solution is to distribute it equally:

$$w^{t+1}_{\alpha,\gamma} = w^t_{\alpha,\gamma} + \frac{s}{x} \tag{4}$$

where s is the sum of the differences of all link values, egressing node α, updated in the previous case,

$$s = \sum_{\beta} (w^{t+1}_{\alpha,\beta} - w^t_{\alpha,\beta})$$

and variable x is the number of neighbours of agent α that did not play.

This policy explores alternative partners if $u < u_P$, since the probability of selecting others in the next game round is increased.

Equation (4) combined with equation (3) are able to achieve the condition expressed by equation (1). If a cooperative agent selects an uncooperative, the corresponding weight will decrease towards zero. The difference is distributed among the weights of players that were not selected. However, the weight of a second uncooperative partner also increases, but not by much. If this second partner is selected, its weight is reduced and distributed among all the partners. The point is that, in the long run, weights of uncooperative agents decrease while weights of cooperative agents will absorb the distributed differences.

3.2 Credit assignment problem

In 2-player games an agent only has one partner. Therefore, the payoff it obtains in a game only depends on its strategy and the strategy of its partner (both fixed, but not necessarily identical). In these games, any player does not have doubts on the quality of its partner. In games with more than two players the situation is different.

In this type of games, there may be partners that the agent should favour instead of others. However, the payoff is not sufficient to establish a differentiated edge update policy. Recall that we have assumed that an agent only knows the payoff it obtains, besides the knowledge of who are its partners. We propose the following procedure:

In the first game, the agent selects $n - 1$ partners randomly, plays the game, but does not apply the update policy. In the following games, the agent randomly replaces one partner and keeps the remaining $n - 2$ partners. It observes the payoff obtained in game t and compares with the payoff obtained in game $t - 1$. Let β_{t-1} be the partner that was selected in game $t - 1$ and β_t the partner that replaced it. Since the agent has only changed one partner, it can compare the payoffs and see which should be favoured. The weight of the link to the agent which provided less payoff is updated using equation (3), weights of kept partners remain unchanged, and weights of remaining neighbours are updated using equation (4).

The drawback of this proposal is that it can only be applied by agent α in games where it selects its partners. The data obtained in games where it is selected by other agents (to play the game), cannot be used by this procedure. In order to use this data we need a more sophisticated agent, that is not analysed in this paper. However, in the next section, we show the usefulness of the simpler approach just described.

4 Experimental Analysis

The purpose of the experiments reported in this paper is to assess the convergence profile of cooperative agents selecting only their equals. To this end, simulations with different proportions and quantities of game strategies were performed.

4.1 The Game

The model presented in the previous section was tested with the Public Good Provision game [7]. The number of players ranged from 3 to 8. For results in 2-player games please refer to [11].

The Public Good Provision game is generally used to model situations where a group of persons has to contribute for a common good [3, 13]. The more people contribute, the greater is the average payoff. However, contribution is costly, so shirking is the rational choice, provided there are at least some players who contribute to the good.

In the Public Good Provision game, an agent that contributes to the good, incurs in a cost c. The good is worth w. Let x be the proportion of agents that provides the good. The payoff of an agent that provides the good is $wx - c$ while agents that defect get wx. The game has a single iteration. The strategy used by agents is probabilistic and is defined by parameter p_p which is the probability to provide the good.

4.2 The Population

Regarding the agents, different strategies were used, which can be roughly classified in how cooperative they are:

S1 A cooperative strategy that always provides the good, $(p_p = 1.0)$.
S2, S3 Two strategies that provide with probabilities 0.7 and 0.3, respectively.
S4 One uncooperative strategy that does not provide the good, $(p_p = 0.0)$.

The number of strategies of each type varied in $\{8, 16, 32\}$. Total population varied in $\{32, 40, 48, \ldots, 128\}$. Moreover, we also performed simulations without strategies **S2** and **S3** thus having only deterministic strategies. This allows us to study convergence for different proportions and quantities of cooperative strategies. Initial edge weight was set to $1/(|\mathbb{P}| - 1)$ so that every player had the same chance of being selected.

Each simulation consisted of 1,000 rounds of games, except for one case where we ran 10,000 rounds, to test convergence. In each round all agents played at least one game, since the following steps were performed per round for every agent: select $n - 1$ partners proportionally to the edge weights, play the game, update the edge weights of the agent that selected partners.

4.3 Results

We have plotted the average probability of agents with strategy $p_p = 1.0$ to select agents from the four strategies. The plots also show the standard deviation. Unless mentioned, the plots were taken at the $1,000^{\text{th}}$ round. Figure 1 shows simulations with only deterministic strategies. Figure 2 shows the results organised by number of cooperative strategies, while figure 3 shows the results organised by number of players in a game. Both latter figures refer to simulations with all four types of strategies.

In all figures we plot an average taken over all possible population sizes and compositions. In figures 1 and 2 we also average results over all game configurations (3 to 8 players).

5 Discussion

In the absence of stochastic strategies, results show that cooperative agents almost always select their equals as partners independently of the conditions. From
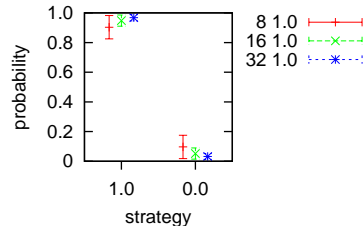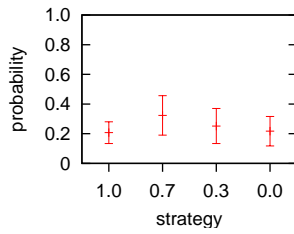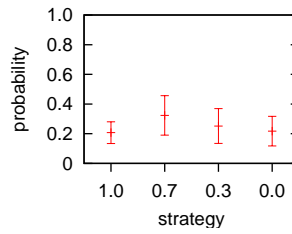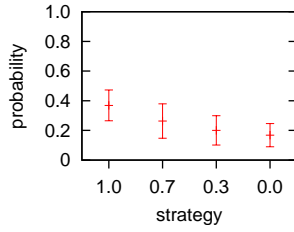
Fig. 1. Results from simulations with only deterministic strategies. Vertical axis represents the probability of strategy (1.0) choosing a strategy in the horizontal axis.
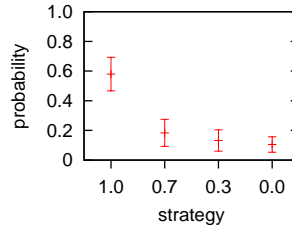

(a) 8 agents with strategy (1.0).

(b) 8 agents with strategy (1.0). Edge weights taken at the $10,000^{\text{th}}$ round.
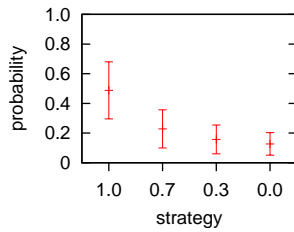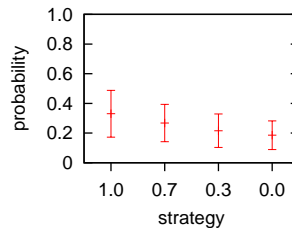
(c) 16 agents with strategy (1.0).

(d) 32 agents with strategy (1.0).

Fig. 2. Simulation results per number of cooperative strategies. Vertical axis represents the probability of strategy (1.0) choosing a strategy in the horizontal axis.

(a) 3-player game.

(b) 8-player game.

Fig. 3. Simulation results per number of players in game. Vertical axis represents the probability of strategy (1.0) choosing a strategy in the horizontal axis.

figure 1 we can observe that the average is high (for choosing (1.0) strategies) while the standard deviation is small.

In the presence of stochastic strategies **S2** and **S3** results change substantially. The number of cooperators influences convergence: the more cooperators, the faster is convergence. We notice higher probability of cooperators selecting their equals for larger numbers of cooperators (compare figures 2(a), 2(c) and 2(d)).

One would expect that stochastic strategies would induce instability in convergence. Suppose we have partners (β_1, β_2) at round t. Both are stochastic and have provided in this round. If β_1 is replaced by a cooperative agent (always provides) and in the next round β_2 does not contribute, then this fact will be imputed to the cooperative agent. However, if we increase the number of rounds to 10,000, the plots obtained are nearly identical as can be seen by comparing figures 2(a) and 2(b), probabilities computed at the $1,000^{\text{th}}$ and $10,000^{\text{th}}$ rounds, respectively. This means that probabilities have converged to a stable situation by iteration 1,000.

The number of players in a game also influences convergence (see figures 3(a) and 3(b) that refer to 3 and 8 players, respectively). The higher is this number, the longer it takes for agents with (1.0) strategy to only select themselves as partners of interaction.

Notice that in spite of averaging a large number of simulations with different parameters (varying population size and composition, and also the number of players, for cases of figures 2 and 3), standard deviation is always small. We can infer that all these parameter values, except the number of cooperative agents, do not influence significantly the results.

6 Future Work

The model we have shown will be analysed in an evolutionary setting where strategies are replaced by some selection policy. Other network structures, besides the panmictic used in this paper, will also be considered.

Stochastic strategies impair convergence of deterministic cooperative strategies, when their number is low. Further analysis of their impact will be carried. The algorithm we have presented in section 3.2 can be modified in order to distinguish deterministic from stochastic strategies. One avenue of researching this is instead of changing one partner per round, keep the same partners when the payoff is equal or higher than u_P, and only change partners when the payoff is lower.

Acknowledgements

References

1. Pat Barclay. Reputational benefits for altruistic punishment. *Evolution of Human Behaviour*, 27:325–344, 2006.
2. Christoly Biely, Klaus Dragosits, and Stefan Thurner. The prisoner's dilemma on co-evolving networks under perfect rationality. *Physica D: Nonlinear Phenomena*, 228(1):40 – 48, 2007.
3. Calvin Blackwell and Michael McKee. Only for my own neighborhood? Preferences and voluntary provision of local and global public goods. *Journal of Economic Behavior & Organization*, 52:115–131, 2003.
4. F.A.C.C. Chalub, F.C. Santos, and J.M. Pacheco. The evolution of norms. *Journal of Theoretical Biology*, 241(2):233–240, 2006.
5. Chrysanthos Dellarocas. Goodwill hunting: An economically efficient online feedback mechanism in environments with variable product quality. In Rino Falcone and Larry Korba, editors, *5th Workshop on Deception, Fraud and Trust in Agent Societies*, pages 26–40, 2002.
6. Franklin M. Fisher. The stability of the cournot oligopoly solution: The effects of speeds of adjustment and increasing marginal costs. *The Review of Economic Studies*, 28(2):125–135, 1961.
7. Herbert Gintis. *Game Theory Evolving - A problem-centered introduction to modeling strategic interaction*. Princeton University Press, 2000.
8. Christoph Hauert. Spatial effects in social dilemmas. *Journal of Theoretical Biology*, 2006.
9. Joseph Henrich and Robert Boyd. Why people punish defectors: Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. *Journal of Theoretical Biology*, 208:79–89, 2001.
10. Leslie Luthi, Mario Giacobini, and Marco Tomassini. Synchronous and asynchronous network evolution in a population of stubborn prisoners. In *CIG*. IEEE, 2005.
11. Ommited. Analysis of a resource sharing game. submitted to EPIA 2009.
12. Francisco C. Santos, Marta D. Santos, and Jorge M. Pacheco. Social diversity promotes the emergence of cooperation in public goods games. *Nature*, 454:213–216, July 2008.
13. Frans van Dijk, Joep Sonnemans, and Frans van Winden. Social ties in a public good experiment. *Journal of Public Economics*, 85:275–299, 2002.