On the Structural Robustness of Evolutionary Models of Cooperation

Segismundo S. Izquierdo[1] and Luis R. Izquierdo[2]

[1] Social Systems Engineering Centre (INSISOC), University of Valladolid, 47011, Spain
[2] The Macaulay Institute, Craigiebuckler, AB15 8QH, Aberdeen, UK
segis@eis.uva.es, l.izquierdo@macaulay.ac.uk

Abstract. This paper studies the structural robustness of evolutionary models of cooperation, i.e. their sensitivity to small structural changes. To do this, we focus on the Prisoner's Dilemma game and on the set of stochastic strategies that are conditioned on the last action of the player's opponent. Strategies such as Tit-For-Tat (TFT) and Always-Defect (ALLD) are particular and classical cases within this framework; here we study their potential appearance and their evolutionary robustness, as well as the impact of small changes in the model parameters on their evolutionary dynamics. Our results show that the type of strategies that are likely to emerge and be sustained in evolutionary contexts is strongly dependent on assumptions that traditionally have been thought to be unimportant or secondary (number of players, mutation-rate, population structure…). We find that ALLD-like strategies tend to be the most successful in most environments, and we also discuss the conditions that favor the appearance of TFT-like strategies and cooperation.

Keywords: Evolution of Cooperation; Evolutionary Game Theory; Iterated Prisoner's Dilemma; Tit for Tat; Agent-based Modeling.

1 Introduction

The evolutionary emergence and stability of cooperation is a problem of fundamental importance that has been studied for decades in a wide range of disciplines. The value of understanding such a question is clear: in the social and biological sciences, the emergence of cooperation is at the heart of subjects as diverse as the first appearance of life, the ecological functioning of countless environmental interactions, the efficient use of natural resources, the development of modern societies, and the sustainable stewardship of our planet. From an engineering point of view, the problem of understanding how cooperation can emerge and be promoted is crucial for the design of efficient decentralized systems where collective action can lead to a common benefit but individual units may (purposely or not) undermine the collective good for their own advantage.

At the most elementary level, the problem of cooperation can be formalized using the symmetric Prisoner's Dilemma (PD), a two-person game where each player can either cooperate or defect. The payoff that players gain when they both cooperate (R)

E. Corchado et al. (Eds.): IDEAL 2006, LNCS 4224, pp. 172 – 182, 2006.

is greater than the payoff obtained when they both defect (P); a single cooperator obtains S, whereas a single defector receives T. The essence of the problem of cooperation is captured by the fact that both players prefer any outcome in which the opponent cooperates to any outcome in which the opponent defects ($T > R > P > S$), but they both have clear incentives to defect. Specifically, both the temptation to cheat ($T > R$) and the fear of being cheated ($S < P$) put cooperation at risk.

Thus, the fundamental challenge of understanding the evolutionary emergence and stability of cooperation can be enlightened, at the most elementary level, by identifying the conditions under which a finite number of units that interact by playing the PD may cooperate. These units might be able to adapt their individual behavior (i.e. learn), or the population of units as a whole may adapt through an evolutionary process (or both). While formalizing the problem of cooperation in this way significantly decreases its complexity (and generality), the question still remains largely unspecified: how many units form the population? How do they interact? What strategies can they use? What is the value of each of the payoffs in the game? and, crucially, what are the processes governing the dynamics of the system?

It is well known since the early years of the study of the evolution of cooperation that, in general, the question of how – if at all – cooperation emerges in a particular system significantly depends on all of the above defining characteristics of the system (see e.g. [1], [2], [3] and [4]). Having recognized this, the method that scientists have naturally followed to advance our formal understanding of the emergence of cooperation has been to study those systems that are tractable with the tools of analysis available at the time. Until not long ago, such tools have derived almost exclusively from the realm of mathematics, and they have given rise to mainstream evolutionary game theory [5]. Mainstream (analytical) Evolutionary Game Theory (EGT) has proven to be tremendously useful, but its use has had important implications in terms of the *classes of systems* that have been investigated, and in terms of the *kind of conclusions* that have been drawn on such systems.

In terms of *classes of systems*, in order to achieve mathematical tractability, EGT has traditionally analyzed *idealized systems*, i.e. systems that *cannot* exist in the real world (e.g. a system where the population is assumed to be infinite). Typically, mainstream EGT has also imposed various other assumptions that simplify the analysis, but which do not necessarily make the system ideal in our terminology (i.e. impossible to exist in the real world). Some examples of such assumptions are: *random* encounters, *infinitely repeated* interactions, *finite* sets of *deterministic* strategies, *proportional fitness rule*, and *arbitrarily small homogenous* invasions. Applying mainstream EGT to non-idealized systems can be very problematic because the validity on non-idealized systems of conclusions drawn from extremely similar idealized systems is not as straightforward as one may think. As an example, Beggs [6] demonstrates that when analyzing some types of evolutionary idealized systems, results can be widely different depending on the order in which certain limits are taken: if one takes the limit as population size becomes (infinitely) large and then considers the limit as the force of selection becomes strong, then one obtains different results from those attained if the order of the limits is inverted. Thus, Beggs [6] warns that "care is therefore needed in the application of these approximations".

The need to achieve mathematical tractability has also influenced the *kind of conclusions* obtained in mainstream EGT. Thus, mainstream EGT has focused on analyzing the stability of incumbent strategies to arbitrarily small mutant invasions, but has not paid much attention to the overall dynamics of the system in terms of e.g. the size of the basins of attraction of different evolutionary stable strategies, or the average fraction of time that the system spends in each of them.

Nowadays it has just become possible to start addressing the limitations of mainstream EGT outlined above. The current availability of vast amounts of computing power through the use of computer grids is enabling us to conduct formal and rigorous analyses of the dynamics of non-idealized systems through an adequate exploration of their sensitivity both to basic parameters and to their structural assumptions. These analyses can complement previous studies by characterizing dynamic aspects of (idealized and non-idealized) systems beyond the limits of mathematical tractability. It is this approach that we follow in this paper.

The specific aim of this paper is to study the structural robustness of evolutionary models of cooperation. To do this, we analyze simple non-idealized models of cooperation and we study their sensitivity to small structural changes (e.g. slight modifications in the way players are paired to play the PD, or in how a generation is created from the preceding one). The impact of the assumptions that we study here has not been, to our knowledge, investigated in a formal and consistent way before arguably because a) it is only recently that we can thoroughly analyze non-idealized models, and/or because b) the effect of such assumptions has been considered unimportant. Thus, in broader terms, our results also shed light on the robustness of the conclusions obtained from EGT as we know it nowadays – i.e. can these conclusions be readily applied to non-idealized systems?

Following this introduction, in section 2 we review some previous work on the robustness of evolutionary models of cooperation. Section 3 describes our modeling framework: EVO-2x2. In section 4 we provide and discuss the main results obtained, and finally, in section 5, we present our conclusions.

2 Previous Work

In this section we report previous work that has shed light on the robustness of evolutionary models of cooperation. We find it useful to place these models in a fuzzy spectrum that goes from mathematically tractable models with strict assumptions that limit their applicability (e.g. work on idealized systems), to models with the opposite characteristics. The rationale behind the construction and use of such a spectrum is that when creating a formal model to investigate a certain question (e.g. the evolution of cooperation), there is often a trade-off between the applicability of the model (determined by how constraining the assumptions embedded in the model are) and the mathematical tractability of its analysis (i.e. how deeply the functioning of the model can be understood given a certain set of available tools of analysis).

The former end is mostly populated by models *designed to* ensure its mathematical tractability. Near this end we find papers that study the impact of some structural assumptions, whilst still keeping others which ensure the model remains tractable and which, unfortunately, also tend to make the model retain its idealized nature. Gotts et al. [4] review many of such papers in sections 2 and 4. Some of these investigations have considered finite vs. infinite populations [7, 8, 9], different pairing settings or population structures (see section 6 in [4]), deterministic vs. stochastic strategies [10, 11], and finite vs. infinitely repeated games [12]. While illuminating, the applicability of most of these studies is somewhat limited since, as mentioned before, the models investigated there tend to retain their idealized nature.

Near the opposite end, we find models that tend to be slightly more applicable (e.g. they consider non-idealized systems), but they are often mathematically intractable. It is from this end that we depart in this paper. To our knowledge, the first relevant study with these characteristics was conducted by Axelrod [13]. Axelrod had previously organized two open tournaments in which the participant strategies played an iterated PD in a round robin fashion [1]. Tit for Tat (TFT) was the winner in both tournaments, and also in an *ecological analysis* that Axelrod [1] conducted after the tournaments. Encouraged by these results, Axelrod [13] investigated the generality of TFT's success by studying the evolution of a randomly generated population of strategies (as opposed to the arguably arbitrary set of strategies submitted to the tournament) using a particular genetic algorithm. The set of possible strategies in this study consisted of all deterministic strategies able to consider the 3 preceding actions by both players. From this study, Axelrod [13] concluded that in the long-term, "reciprocators […] spread in the population, resulting in more and more cooperation and greater and greater effectiveness". However, the generality of Axelrod's study [13] is doubtful for two reasons: (1) he used a very specific set of assumptions, the impact of which was not tested, and (2) even if we constrain the scope of his conclusions to his particular model, the results should not be trusted since Axelrod only conducted 10 runs of 50 generations each. As a matter of fact, Binmore [14, 15] cites unpublished work by Probst [16] that contradicts Axelrod's results.

In a more comprehensive fashion, Linster [17] studied the evolution of strategies that can be implemented by two-state Moore machines in the infinitely repeated PD. He found a strategy called GRIM remarkably successful. In particular, GRIM was significantly more successful than TFT. GRIM always cooperates until the opponent defects, in which case it switches to defection forever. Linster [17] attributed the success of GRIM over TFT to the fact that GRIM is able to exploit poor strategies while TFT is not. Linster's investigation was truly remarkable at its time, but technology has advanced considerably since then, and we are now in a position to expand his work significantly by conducting parameter explorations beyond what was possible before. As an example, note that Linster [17] could only consider deterministic strategies and one specific value for the mutation rate; furthermore, in the cases he studied where the dynamics were not deterministic, there is no guarantee that his simulations had reached their asymptotic behavior.

In the following section we describe the relevant aspects of our modeling framework, which is aimed at facilitating a more consistent and systematic exploration of the impact of competing assumptions in non-idealized evolutionary models of cooperation.

3 Our Modeling Framework: EVO-2x2

EVO-2x2 was developed using NetLogo [18]. In EVO-2x2 there is a population of ***num-players*** players. Events occur in discrete time-steps, which can be interpreted as successive generations. At the beginning of every generation every player's payoff (which denotes the player's fitness) is set to zero. Then, every player is paired with another player to play a 2-player match, according to one of two possible pairing algorithms (***pairing-settings***):

- *random pairings*: Pairs are made at random, without any bias.
- *children together*: Players are paired preferentially with their siblings (and at random among siblings). Once all the possible pairs between siblings have been made, the rest of the players are paired at random. This procedure was implemented because it seems plausible in many biological contexts that individuals belonging to the same family tend to interact more often among them than with individuals from other families.

Every player plays one single match per generation. Each match consists of a number of sequential rounds (***rounds-per-match***). In each round, the two members of the pair play a symmetric PD once. The action selected by each of the players determines the magnitude of the payoff that each of them receives in that round (***CC-payoff***, ***CD-payoff***, ***DC-payoff***, ***DD-payoff***). The total payoff that a player obtains in a match is the sum of the payoffs obtained in each of the rounds. Players differ in the way they play the match, i.e. they generally have different strategies. The strategy of a player is determined by three numbers between 0 and 1:

- *PC*: Probability to cooperate in the first round.
- *PC/C*: Probability to cooperate in round n ($n > 1$) given that the other player has cooperated in round ($n - 1$).
- *PC/D*: Probability to cooperate in round n ($n > 1$) given that the other player has defected in round ($n - 1$).

The set of possible values that *PC*, *PC/C* or *PC/D* can take depends on the value of the binary variable ***infinite-strategies?***, which is either on or off. If on (default option), the set of possible values is any (floating-point) number between 0 and 1. If off, only ***num-strategies*** (≥ 2) values are allowed for each of the variables *PC*, *PC/C*, and *PC/D*; the permitted values are evenly distributed between 0 and 1 (both included). Once every player has played one – and only one – match, two evolutionary processes come into play to replace the old generation with a brand new one: natural selection (***selection-mechanism***) and mutation (***mutation-rate***). Successful players (those with higher payoffs) tend to have more offspring than unsuccessful ones. This marks the end of a generation and the beginning of a new one, and thus the cycle is completed. In this paper we only consider a ***selection-mechanism*** called *roulette wheel*, which involves conducting ***num-players*** replications that form the new generation. In each replication, players from the old generation are given a probability of being chosen to be replicated that is proportional to their total payoff. A mutant is a player whose strategy (the 3-tuple formed by *PC*, *PC/C*, and *PC/D*) has been determined at random. The probability that any newly created player is a mutant is ***mutation-rate***.

4 Results and Discussion

We use EVO-2x2 to conduct a systematic exploration of the parameter space for the PD, in order to assess the impact of various competing assumptions. All the simulations reported in this paper have been run on computer grids.

Defining a state of the system as a certain particularization of every player's strategy, it can be shown that all simulations in EVO-2x2 with positive mutation rates can be formulated as irreducible positive recurrent and aperiodic (sometimes called ergodic) discrete-time Markov chains. Thus, there is a unique long-run distribution over the possible states of the system, *i.e.* initial conditions are immaterial in the long-run [19, Theorem 3.15]. Although calculating such (dynamic) distributions analytically is unfeasible, we can estimate them using the computer simulations. The problem is to make sure that a certain simulation has run for long enough, so the limiting distribution has been satisfactorily approximated. To make sure that this is the case, for each possible combination of parameters considered, we ran 8 different simulations starting from widely different initial conditions. These are the 8 possible initial populations where every individual has the same pure strategy (the 8 corners of the strategy space). Then, every simulation run is conducted for 1,000,000 generations. Thus, in those cases where the 8 distributions are similar, we have great confidence that they are showing a distribution close to the limiting distribution.

A useful summary of the results produced in a simulation run is the accumulated frequency of different types of strategies throughout the course of a simulation run. This is something that can be plotted in a 3D contour plot, and in complementary 2D density plots, as shown in figure 1.

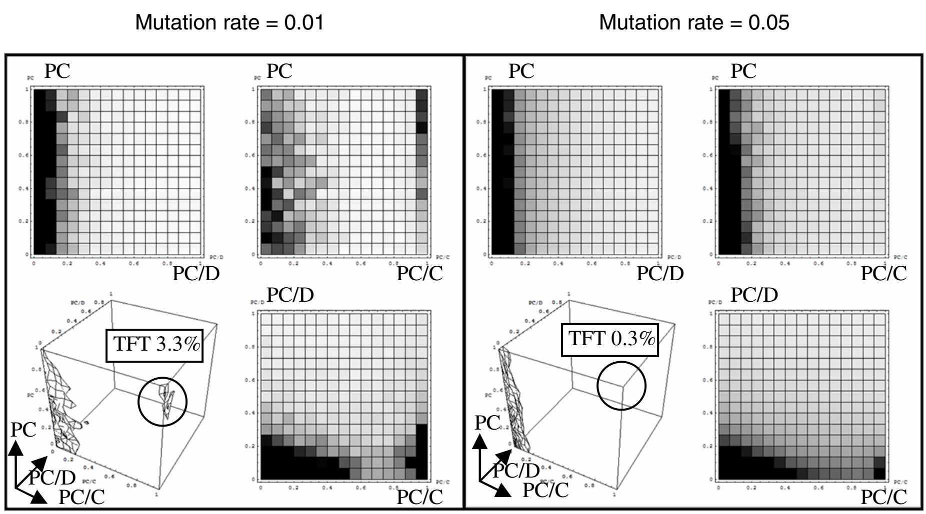

Fig. 1. Influence of the mutation rate on the dynamics of the system. TFT measures the average time that strategies with $PC \geq (13/15)$, $PC/C \geq (13/15)$ and $PC/D \leq (2/15)$ were observed.

Here we report several cases where it can be clearly seen that some of the assumptions that are sometimes thought to have little significance (e.g. mutation-rate, number of players, or population structure) can have a major impact on the type of strategies that emerge and are sustained throughout generations. The payoffs used in all simulations are: ***CC-payoff*** = 3; ***CD-payoff*** = 0; ***DC-payoff*** = 5; ***DD-payoff*** = 1.

The two distributions in figure 1 only differ in the value of the mutation rate used (0.01 on the left, and 0.05 on the right). The distribution on the left shows the evolutionary emergence and (dynamic) permanence of strategies similar to TFT (*PC* ≈ 1, *PC/C* ≈ 1, and *PC/D* ≈ 0). Such strategies do not appear for slightly higher mutation rates (distribution on the right). The other parameter values used were ***num-players*** = 100; ***pairing-settings*** = *random pairings*; ***rounds-per-match*** = 50.

The two distributions in figure 2 only differ in the number of players in the population (100 on the left, and 10 on the right). The distribution on the left shows the evolutionary emergence and (dynamic) permanence of strategies similar to TFT, whereas such strategies do not appear in smaller populations. The other parameter values are: ***pairing-settings*** = *random pairings*; ***rounds-per-match*** = 50; ***mutation-rate*** = 0.01.

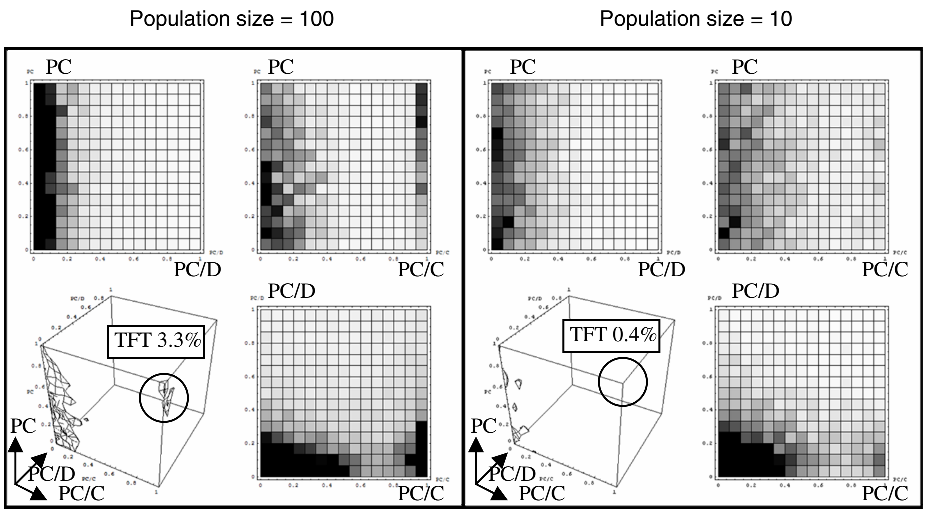

Fig. 2. Influence of the number of players in the population. TFT measures the average time that strategies with *PC* ≥ (13/15), *PC/C* ≥ (13/15) and *PC/D* ≤ (2/15) were observed.

The two distributions in figure 3 only differ in the algorithm used to form the pairs of players (*random pairings* on the left, and *children together* on the right). On the left, strategies tend to be strongly uncooperative, while the distribution on the right is concentrated around strategies similar to TFT. The other parameter values used were: ***num-players*** = 100; ***rounds-per-match*** = 5; ***mutation-rate*** = 0.05.

The two distributions in figure 4 only differ in the set of possible values that *PC*, *PC/C* or *PC/D* can take. For the distribution on the left the set of possible values is any (floating-point) number between 0 and 1, and the strategies are mainly uncooperative, similar to ALLD (*PC* ≈ 0, *PC/C* ≈ 0, and *PC/D* ≈ 0). For the distribution on the right, the set of possible values is only {0, 1}, and the distribution is concentrated in TFT. The other parameter values used were: ***num-players*** = 100; ***mutation-rate*** = 0.05; ***rounds-per-match*** = 10; ***pairing-settings*** **=** *random pairings*.

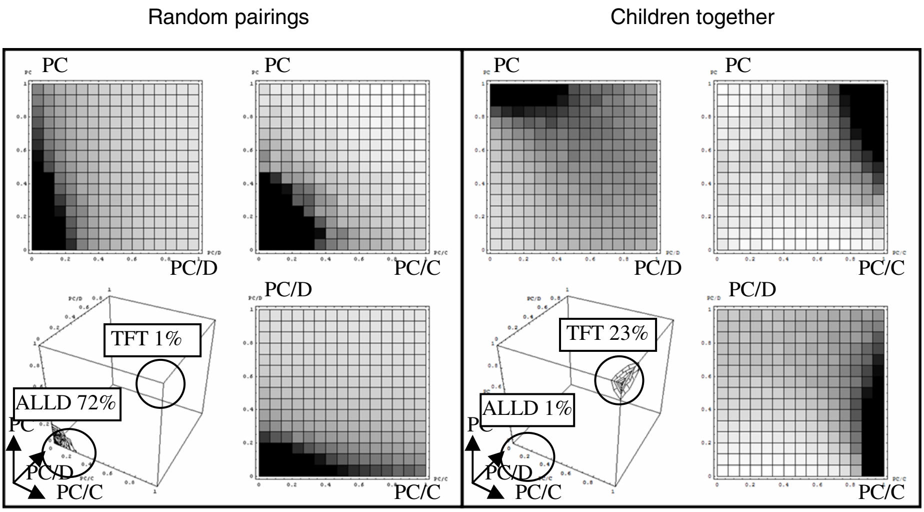

Fig. 3. Influence of different pairing mechanisms. TFT measures the average time that strategies with $PC \geq (10/15)$, $PC/C \geq (10/15)$ and $PC/D \leq (5/15)$ were observed; ALLD measures the average time that strategies with $PC \leq (5/15)$, $PC/C \leq (5/15)$ and $PC/D \leq (5/15)$ were observed.

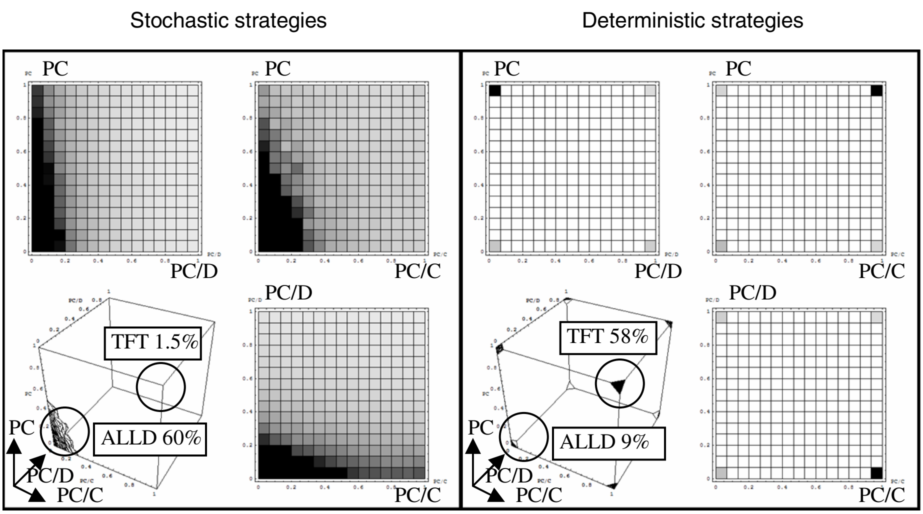

Fig. 4. Stochastic (mixed) strategies vs. deterministic (pure) strategies: influence in the system dynamics. TFT measures the average time that strategies with $PC \geq (10/15)$, $PC/C \geq (10/15)$ and $PC/D \leq (5/15)$ were observed; ALLD measures the average time that strategies with $PC \leq (5/15)$, $PC/C \leq (5/15)$ and $PC/D \leq (5/15)$ were observed.

In figures 5 and 6 we show the effect of gradually increasing the set of possible values for *PC*, *PC/C* and *PC/D* (i.e. ***num-strategies***). Figure 5 shows the (average) number of each possible outcome of the game (CC, CD/DC or DD) in observed series of 10^6 matches (this number of matches is selected so the effect of changing the initial state is negligible, i.e. results are close to the stationary limiting distribution).

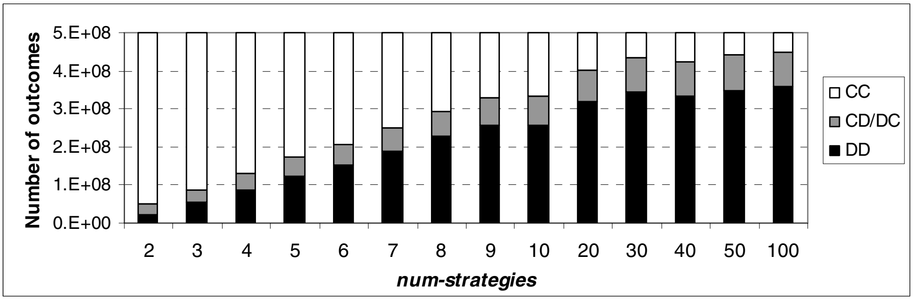

Fig. 5. Influence in the distribution of outcomes (CC, CD/DC or DD) of augmenting the set of possible values for *PC*, *PC/C* and *PC/D*

Figure 6 shows the average values of *PC*, *PC/C* and *PC/D* observed in the same series. Augmenting the set of possible values for *PC*, *PC/C* and *PC/D* undermines cooperation and favors the emergence of ALLD-like strategies. The other parameter values used were: ***num-players*** = 100; ***mutation-rate*** = 0.01; ***rounds-per-match*** = 10; ***pairing-settings*** = *random pairings*.

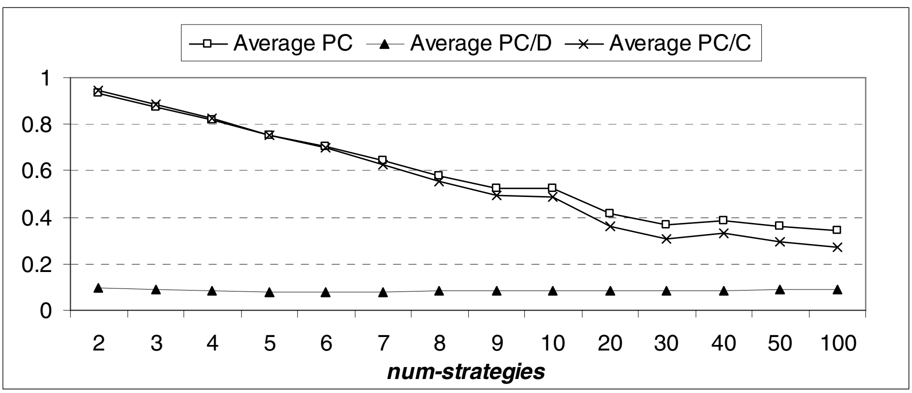

Fig. 6. Influence of augmenting the set of possible values for *PC*, *PC/C* and *PC/D* in the average values of these variables in the population

5 Conclusions

In this paper we have presented, for the Prisoner's Dilemma, several results on the evolutionary dynamics of stochastic strategies that are conditioned on the last action of the player's opponent. Popular strategies such as TFT and ALLD are particular (extreme) cases within this framework, and we have studied the possible appearance and evolutionary robustness of such strategies. Our results show that:

- The type of strategies that are likely to emerge and be sustained in evolutionary contexts is strongly dependent on assumptions that traditionally have been thought to be unimportant or secondary (value of mutation-rate, number of players, population structure…)
- Strategies similar to ALLD tend to be the most successful in most environments.
- Strategies similar to TFT tend to spread best with the following factors: in large populations, where individuals with similar strategies tend to interact more frequently, when only deterministic strategies are allowed, with low mutation rates, and when interactions consist of many rounds.

Acknowledgements

We would like to gratefully acknowledge financial support from the Spanish Ministerio de Educación y Ciencia research project DPI2005-05676 "SIGAME" and from the Scottish Executive Environment and Rural Affairs Department.

References

1. Axelrod, R.: The Evolution of Cooperation. Basic Books USA (1984)
2. Bendor, J., Swistak, P.: 1995. Types of evolutionary stability and the problem of cooperation. Proceedings of the National Academy of Sciences USA **92** (1995) 3596-3600
3. Bendor, J., Swistak, P.: Evolutionary Equilibria: Characterization Theorems and Their Implications. Theory and Decision **45** (1998) 99-159.
4. Gotts, N.M., Polhill, J.G., Law, A.N.R.: Agent-based simulation in the study of social dilemmas. Artificial Intelligence Review, **19**(1) (2003) 3-92.
5. Weibull, J. W.: Evolutionary Game Theory. MIT Press (1995)
6. Beggs, A.: Stochastic evolution with slow learning. Economic Theory **19** (2002) 379-405
7. Imhof, L.A., Fudenberg, D., Nowak M.A.: Evolutionary cycles of cooperation and defection. Proceedings of the National Academy of Sciences USA, **102**(31) (2005) 10797-10800
8. Nowak, M.A., Sasaki, A., Taylor, C., Fudenberg, D.: Emergence of cooperation and evolutionary stability in finite populations. Nature **428** (2004) 646-650
9. Taylor, C., Fudenberg, D., Sasaki, A., Nowak, M.A.: Evolutionary Game Dynamics in Finite Populations. Bulletin of Mathematical Biology **66** (2004) 1621–1644
10. Nowak, M.A., Sigmund K.: The evolution of stochastic strategies in the Prisoners' Dilemma. Acta Applicandae Mathematicae **20** (1990) 247-265
11. Nowak, M.A., Sigmund K.: Tit for tat in heterogeneous populations. Nature **355** (1992) 250-253
12. Nowak, M.A., Sigmund K.: Invasion Dynamics of the Finitely Repeated Prisoner's Dilemma. Games and Economic Behavior **11**(2) (1995) 364-390
13. Axelrod, R.: The Evolution of Strategies in the Iterated Prisoner's Dilemma. In: Davis, L. (Ed.): Genetic Algorithms and Simulated Annealing. London Pitman, Los Altos CA, Morgan Kaufman (1987) 32-41. Reprinted in Axelrod, R.: The complexity of cooperation. Agent-based models of competition and collaboration. Princeton NJ, Princeton University Press (1997)

14. Binmore, K.: Playing Fair: Game Theory and the Social Contract I. MIT Press, Cambridge MA (1994)
15. Binmore, K.: Review of the book: The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration, by Axelrod, R., Princeton NJ, Princeton University Press, (1997). Journal of Artificial Societies and Social Simulation **1**(1) (1998) http://jasss.soc.surrey.ac.uk/1/1/review1.html
16. Probst, D.: On Evolution and Learning in Games. PhD thesis, University of Bonn (1996)
17. Linster, B.: Evolutionary stability in the repeated Prisoners' Dilemma played by two-state Moore machines. Southern Economic Journal **58** (1992) 880-903
18. Wilensky, U.: NetLogo. http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL (1999)
19. Kulkarni, V.G.: Modelling and Analysis of Stochastic Systems. Chapman & Hall/CRC (1995)