A Model-to-Model Analysis of The Repeated Prisoners' Dilemma: Genetic Algorithms *vs.* Evolutionary Dynamics[*]

Xavier Vilà[†]

Universitat Autònoma de Barcelona

June 11, 2008

Abstract

We study the properties of the well known Replicator Dynamics when applied to a finitely repeated version of the Prisoners' Dilemma game. We characterize the behavior of such dynamics under strongly simplifying assumptions (i.e. only 3 strategies are available) and show that the basin of attraction of defection shrinks as the number of repetitions increases. After discussing the difficulties involved in trying to relax the "strongly simplifying assumptions" above, we approach the same model by means of simulations based on genetic algorithms. The resulting simulations describe a behavior of the system very close to the one predicted by the replicator dynamics without imposing any of the assumptions of the analytical model. Our main conclusion is that analytical and computational models are good complements for research in social sciences. Indeed, while on the one hand computational models are extremely useful to extend the scope of the analysis to complex scenarios hard to analyze mathematically, on the other hand formal models can be extremely useful to verify and to explain the outcomes of computational models.

Keywords: Agent-Based Computational Economics, Evolutionary Game Theory, Replicator Dynamics, Model-to-Model Analysis, Repeated Prisoners' Dilemma.
JEL Classification: C63, C72, D82.

1 Introduction

In the growing field of Agent-Based computer simulations applied to social sciences, model replication is considered a key issue. Indeed, asserting whether the observed

[†]Departament d'Economia i d'Història Econòmica. Universitat Autònoma de Barcelona. 08913-Bellaterra (Barcelona). Spain. `Xavier.Vila@uab.cat`

results of a particular simulation of a model are correct or generalizable is a difficult task when no formal (i.e. mathematical) proof is provided. Only replication, comparison, alignment, and other related techniques can shed some light on the validity of simulations. This work contains one such comparison. We put side-by-side two different analysis (mathematical and computational) of the same model: the evolution of strategies in the repeated prisoners' dilemma.

First, we consider the case in which the evolutionary system can be described by a deterministic dynamic system that uses expected values. Using strong simplifying assumptions we are able to solve this case and to produce a complete description of how the process behaves. We also discuss the problems involved when we try to relax some of the assumptions made.

The second approach is a computational simulation in which finite automata are used to represent the strategies played and a decentralized adaptive process based on the models of genetic algorithms to simulate the stochastic evolutionary process. With this technique we can relax some of the strong assumptions used in the first approach and still obtain the same basic results.

We like to think that the limitations of the first approach (analytical) provide a good motivation for the second approach (Agent-Based simulations). Indeed, although both approaches address the same problem, we show that the use of Agent-Based computational techniques allows us to relax hypothesis and overcome the limitations of the analytical approach. On the other hand, it is shown that the analytical model is extremely useful in order to explain the behavior and the results of the computational model.

The choice of the repeated prisoners' dilemma to conduct the experiment described above is not arbitrary. It is a well know and largely studied game, and many things about it have been learned thanks to the tools of formal game theory. But when the game is studied from an evolutionary perspective, the results are not always clear. The works by Boyd and Lorberbaum (1987) and Binmore and Samuelson (1992), for instance, show that evolutionary stable solutions may fail to exists in many versions of the game. Experiments and simulations, on the other hand, like the ones conducted by Axelord (1984), Nowak and Sigmund (1992), or Miller (1996), seem to suggest that Tit-for-tat (and other similar strategies) prevail in most of the situations. Thus, the interest of our resarch is putting these two approaches, analytical and computational, side-by-side to achieve a better understanding of the evolutionary behavior of players in the repeated prisoners' dilemma

2   The Analytical Model

The basic stage game (Prisoners' Dilemma) that players will play repeatedly is given by

	C	D
C	3,3	0,5
D	5,0	1,1

2

We now consider the repeated version of the game played a finite number of rounds R. In order to keep things simple, we only consider three possible strategies (as in Nowak et al. (2004)):

- D: Always defect

- C: Always cooperate

- T: Tit-for-Tat

as they are the three strategies that have deserved a higher attention in almost all the literature dealing with the Repeated Prisoner's Dilemma from an evolutionary point of view. Given the above, the repeated game can be represented as follows

	C	D	T
C	$3R,3R$	$0,5R$	$3R,3R$
D	$5R,0$	R,R	$5+(R-1),(R-1)$
T	$3R,3R$	$(R-1),5+(R-1)$	$3R,3R$

Thus, for instance, when a D-type strategy meets a T-type strategy, the former gets 5 in the first round and then 1in each subsequent round ($5+(R-1)$ in total), while the later gets 0 first and then 1in each subsequent round ($R-1$ in total).

2.1 The Replicator Dynamics analysis

Let $p_t(C)$ be the probability that, at time t, a player in this population is an "always cooperate" type, and the same for $p_t(D)$ and $p_t(T)$. We thus have that $p_t(C)+p_t(D)+p_t(T)=1 \ \forall t$.

The replicator dynamics states that the rate of change of such probabilities is a function of the relative performance of each strategy with respect to the average performance of the population. In this sense, given $p_t(C), p_t(D), p_t(T)$, the expected payoff for each strategy is:

$$E_t\pi(C) = 3Rp_t(C)+0p_t(D)+3Rp_t(T) = 3R(P_t(C)+p_t(T))$$

$$E_t\pi(D) = 5Rp_t(C)+Rp_t(D)+(5+(R-1))p_t(T)$$

$$E_t\pi(T) = 3Rp_t(C)+(R-1)p_t(D)+3Rp_t(T) = 3R(P_t(C)+p_t(T))+(R-1)p_t(D)$$

and the average payoff will be:

$$E_t\bar{\pi} = E_t\pi(C)p_t(C)+E_t\pi(D)p_t(D)+E_t\pi(T)p_t(T)$$

Notice that since $p_t(C)+p_t(D)+p_t(T)=1 \ \forall t$ only two dimensions matter. Hence, the replicator dynamics in this case is given by:

$$\frac{\partial p_t(C)}{\partial t} = p_t(C)(E_t\pi(C)-E_t\bar{\pi})$$

3

$$\frac{\partial p_t(D)}{\partial t} = p_t(D)(E_t\pi(D) - E_t\bar{\pi})$$

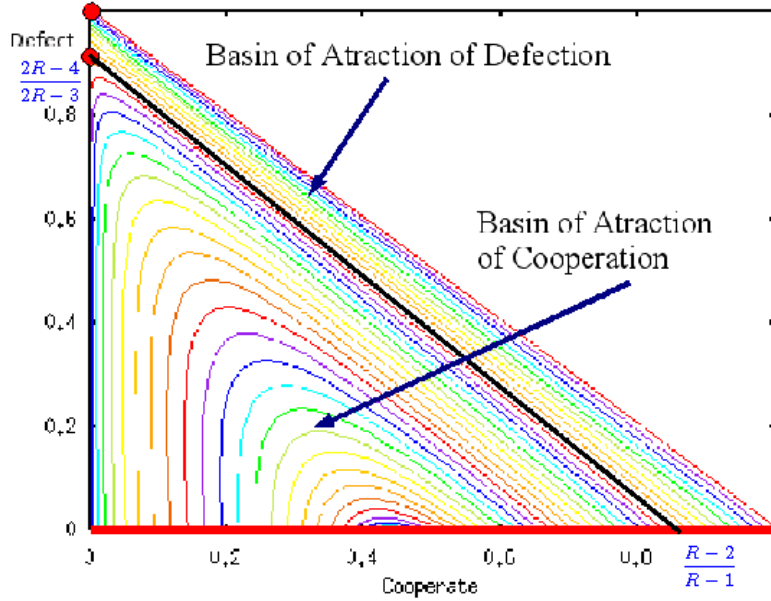and the corresponding vector field showing the trajectories of the system is given in Figure 1

Figure 1: Vector Field

The horizontal and vertical axis in Figure 1 correspond to $p_t(C)$ and $p_t(D)$ respectively. Thus, the three vertexes of the triangle $((1,0), (0,1), \text{and} (0.0))$, correspond to the states $p_t(C) = 1$, $p_t(D) = 1$, and $p_t(T) = 1$ respectively. The trajectories that represent the evolution of the system are divided in two areas or basins of attraction, one for Defection and another for Cooperation. Stationary points of the system are marked red: $(0,1)$ corresponding to everybody playing *always defect* $p_t(D) = 1$, $(0, \frac{2R-4}{2R-3})$, and all the points in the line that goes from $(0,0)$ to $(1,0)$ that corresponds to points with no *defectants*, that is, $p_t(D) = 0$ and $p_t(C) + p_t(T) = 1$. Only the point $(0,1)$ corresponding to $p_t(D) = 1$ is asymptotically stable in the sense that if the system is slightly perturbed away from $(0,1)$, any trajectory will bring it back to the same point. The singular point $(0, \frac{2R-4}{2R-3})$ can only be reached if the system starts somewhere in the line that goes from $(0, \frac{2R-4}{2R-3})$ to $(\frac{R-2}{R-1}, 0)$, which occurs with zero probability.

An important result is that the relative size of these basins of attraction depends on the number of repetition R. That is, if the system starts at random, the probability of

reaching the point $(0,1)$ (everybody defecting) or the line $(0,0) \rightarrow (1,0)$ (everybody cooperating) depends on R. Thus, we can compute what is the *expected per-round payoff* a priori depending on R.

$$E\bar{\pi} = (\frac{2R-4}{2R-3})(\frac{R-2}{R-1}) \cdot 3 + (1 - (\frac{2R-4}{2R-3})(\frac{R-2}{R-1})) \cdot 1$$

Figure 2 shows the behavior of the *expected per-round payoff* as a function of R. We observe that it grows rapidly as the number of repetitions (R) increases. In fact, $E\bar{\pi} \rightarrow 3$ as $R \rightarrow \infty$

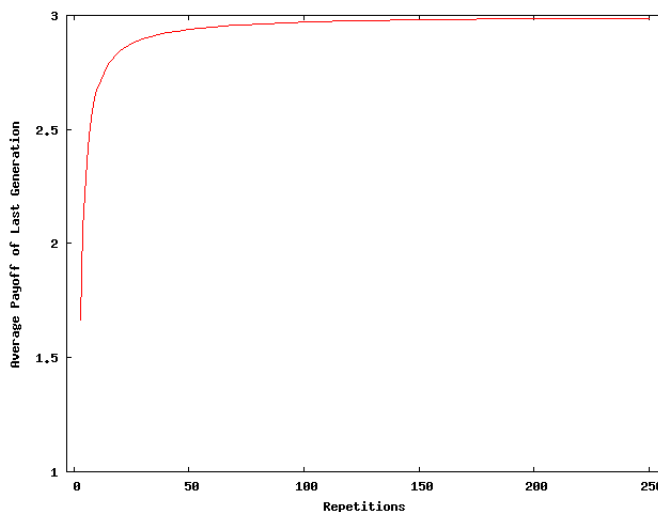Figure 2: Expected payoff as a function of R

3   The Computational Model

Given the analysis above, the dynamics seem to suggest that there is room for cooperation. At least for a broad range of initial conditions, the trajectories lead to some point in the horizontal axis corresponding to a population consisting of *only C* and *T* strategies.

Nevertheless, such analysis is extremely partial since we are only considering 3 strategies at a time, namely C, D, and T. One can easily see that extending this approach to a more general case (with more strategies considered) is a difficult task.

To overcome this limitation, we develop a computer simulation in which the strategies are represented by finite automata of size four and a Genetic Algorithm routine is used to simulate the evolutionary process as in Miller (1996). In most of the cases, the results of such simulations produce the outcome in Figure 3, in which the evolution of the (per round) average payoff is displayed.
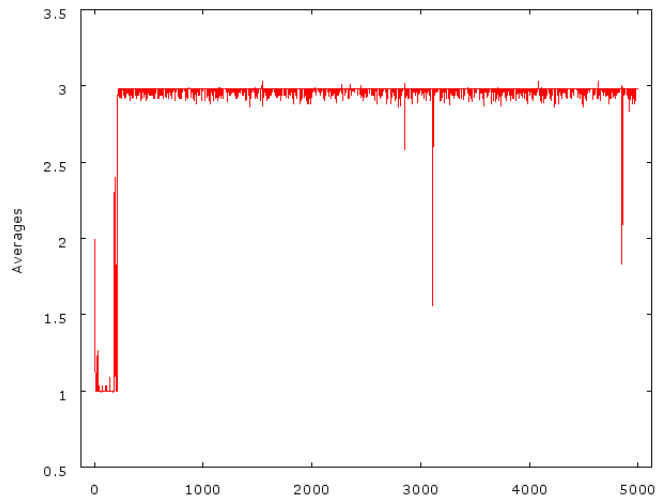
5

Figure 3: Evolution of the average per-round payoff when Cooperation is the result

Because the final average payoff is 3 we can conclude that all players follow a cooperative strategy.

In other cases, though, cooperation is not the final result as the evolution of the average payoff results as in Figure 4.
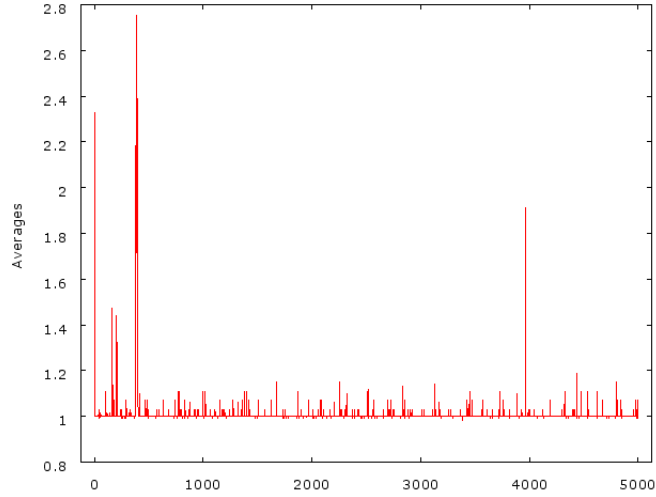

Figure 4: Evolution of the average per-round payoff when Defection is the result

In both cases, though, the resemblance between the vector field in Figure 1 and the evolution of payoffs in Figures 3 and 4 is very appealing. When the final result is

cooperation (as in Figure 3), both in the replicator dynamic analysis and in the simulations, the evolutionary process seems to favor the growth of Defectant strategies at first, and then these disappear and Cooperative strategies start to replicate to end up with the payoff corresponding to the cooperative behavior. On the contrary, when the final result is *defection,* the evolution goes "monotonically" towards that point.

How often each of these two results occurs ? Given that in the analytical model we have found out that the answer to this question depends upon the number of repetitions R, we check whether R also has an effect in the computational model. In this sense, Figure 5 complements Figure 2 by showing how the final observed average payoff of the simulations[1] (the payoff of generation 5000) depends on R. For robustness, we do such exercise with two different crossover rules and with no crossover
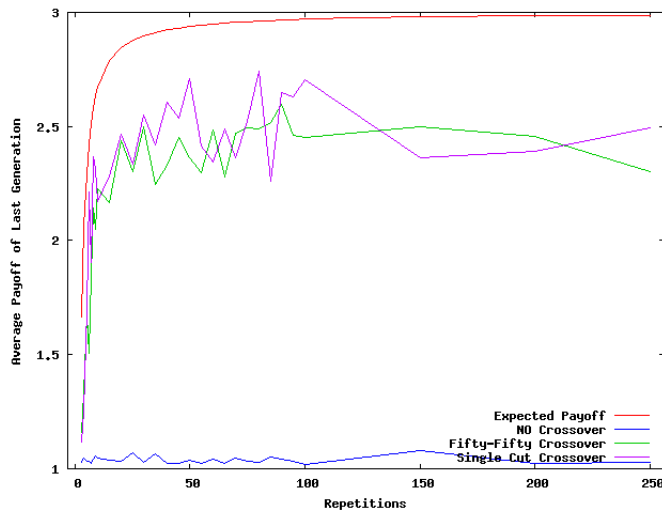

Figure 5: Expected and observed payoffs as a function of R

Figure 5 shows how the behavior of the simulations also resembles what we obtained theoretically in the previous section. That is, as the number of repetitions R grows, the higher is the probability of reaching cooperation as the final result and hence, the higher is the average payoff (both theoretical and empirically).

In this sense, it seems that the use of Genetic Algorithms to simulate the evolutionary process closely matches the behavior predicted by the replicator dynamics while avoiding the strong limitation of considering only 3 possible strategies.

4   Conclusions

We have studied the evolution of strategies in the well known Repeated Prisoner's Dilemma using two different approaches: one analytical based on the replicator dy-

---

[1]For each value of R we run 100 simulations and compute the average payoff of the last generation (5000)

namics and one computational based on genetic algorithms. We show that the results obtained from the two approaches coincide almost completely in the sense that,

1. The two approaches produce the same two possible outcomes: evolution towards defection or evolution towards cooperation.

2. In the two approaches, the path towards the equilibrium are similar: monotonic when going towards defection, non-monotonic when going towards cooperation

3. In the two approaches, the proportion of times each of the two possible results occurs is also similar and depends on the number of repetitions of the game.

The scope of these conclusions, though, might be somehow limited. The reason is that, while in the analytical model only three strategies are considered (Always cooperate, Always defect, and Tit-for-tat), the computational model deals with finite automata of size 4, which can represent a very large number of different strategies. Nevertheless, one generally observes that from a starting random population of strategies (represented by finite automata), the genetic algorithm rapidly reduces the number of "working" strategies and, at the end, only strategies similar to the three used in the analytical model appear. Also, in Vilà (2008) we discuss other genetic algorithm operators that can deal with this issue, and the results are not different from the ones presented here.

The results of this research seem to suggest that, in our opinion, analytical and computational models are good complements for resrach in social sciences. Indeed, while on the one hand computational models are extremely useful to extend the scope of the analysis to complex scenarios hard to analyze mathematically, on the other hand formal models can be extremely useful to verify and to explain the outcomes of computational models.

References

[1] Axelrod, R. and Hamilton, W.: The Evolution of Cooperation. Science **211** (1981) 1390-1396

[2] Axelrod, R.: The Evolution of Cooperation. New York: Basic Books (1984)

[3] Miller, J. H.: The coevolution of automata in the repeated Prisoner's Dilemma. Journal of Economic Behavior and Organization **29** (1996) 87–112

[4] Binmore, K. and Samuelson, L: Evolutionary Stability in Repeated Games Played by Finite Automata. Journal of Economic Theory **57** (1992) 278-305

[5] Boyd, R. and Lorberbaum J.: No Pure Strategy is Evolutionarily Stable in the repeated Prisoner's Dilemma Game. Nature **327** (1987) 58-59

[6] Nowak, M. A., Sasaki, A., Taylor, C., and Fudenberg, D.: Emergence of cooperation and evolutionary stability in finite populations. Nature **428** (2004) 646 – 650

[7] Nowak, M. and Sigmund, K: Tit for Tat in Heterogeneous Populations. Nature **355** (1992) 250-253

[8] Vilà, X.: A Model-To-Model Analysis of Bertrand Competition. Journal of Artificial Societies and Social Simulation **11(2)** (2008) <http://jasss.soc.surrey.ac.uk/11/2/11.html>.