Studies on Dynamics in the Classical Iterated Prisoner's Dilemma with Few Strategies

Is There Any Chaos in the Pure Dilemma ?

Philippe Mathieu, Bruno Beaufils, and Jean-Paul Delahaye

Laboratoire d'Informatique Fondamentale de Lille (UPRESA 8022 CNRS)
Université des Sciences et Technologies de Lille – UFR d'IEEA – Bât. M3
59655 Villeneuve d'Ascq Cedex – FRANCE
`mathieu@lifl.fr`, `beaufils@lifl.fr`, `delahaye@lifl.fr`

Abstract. In this paper we study Classical Iterated Prisoner's Dilemma (CIPD) dynamics of pure strategies in a *discrete* and *determinist* simulation context. We show that, in some very rare cases, they are not quiet and ordered. We propose a classification of ecological evolutions into categories which represent complex dynamics, such as oscillatory movements. We also show that those *simulations* are very sensitive to initial conditions. These experimentations could call into question classical conclusions about interest of cooperation between entities playing CIPD. They may be used to explain why it is not true that cooperation is always the convergent phenomenon observed in life.

1 The Classical Iterated Prisoner's Dilemma

When they were at the RAND Corp. Merill M. Flood and Melvin Dresher tried to introduce some kind of *irrationality* in Game Theory, [13, 11]. They introduced a simple two person, non zero-sum, non cooperative and simultaneous game, [6]. This game, very simple to describe, covers a large scale of *real life* situations and seems to catch the definition of conflicts of interests. Thus a lot of different kind of work has been done on it, involving not only mathematicians, but also social, zoological, biological as well as computer scientists. The game becomes the most used theoretical model for studying the cooperation and the evolution of cooperation in population of agents.

The game, called the Prisoner's Dilemma, could be described very simply in the following way: let us meet two *artificial agents* having two choices (two *strategies*):

- COOPERATE, let us write `C`, and say to be *nice*
- DEFECT, let us write `D`, and say to be *naughty*

The payoff of each player depends on the moves played by the two agents. Tab. 1 names the score of each case.

The dilemma comes when exploitation of one by the other (T) is better payed than cooperation between the two (R), which itself pays more than a case where

Table 1. CIPD payoff matrix. Row player score are given first.

	Cooperate	Defect
Cooperate	$R = 3$, $R = 3$ *Reward* for mutual cooperation	$S = 0$, $T = 5$ *Sucker*'s payoff *Temptation* to defect
Defect	$T = 5$, $S = 0$ *Temptation* to defect *Sucker*'s payoff	$P = 1$, $P = 1$ *Punishment* for mutual defection

the two tried two exploite each other (P), which finally is a better choice than to be exploited (S). This can be formalised as:

$$S < P < R < T \tag{1}$$

The dilemma stands on the fact that individual interest (Nash equilibrium) differs from collective one (Pareto issues).
The one shot game, involving rational agents and pure strategies, is solved in Game Theory by the Nash equilibrium, which is to always betray its partner: choosing the `D` strategy. In an iterated version players meet each other more than one time. The payoff of an agent is then simply the sum of each of its meeting's payoff. The game is called the Classical Iterated Prisoner's Dilemma (CIPD).
In order to favour cooperation over defection the following constraint is added:

$$S + T < 2R \tag{2}$$

A classical choice of payoff values, mainly introduced by [1], is given in Tab. 1. With such an iterated game what the opponent did on previous moves clearly influences the way an agent will play on next ones. It is then possible to define more strategies than with the one shot version. Let us define some simple ones, some of which will be used in our experimentations :

`all_c` corresponds to the `C` strategy of the one shot game: it always plays `C`
`all_d` corresponds to the `D` strategy of the one shot game: it always plays `D`
`tit_for_tat` cooperates on the first move then plays opponent's previous move.
`per_cd` plays periodically `C` then `D`, let us note (`CD`)$*$
`per_ddc` plays (`DDC`)$*$
`per_ccd` plays (`CCD`)$*$
`per_ccccd` plays (`CCCCD`)$*$
`soft_majo` plays opponent's majority move, cooperating in case of equality
`prober` plays (`DCC`), then it defects in all other move if opponent has cooperated in move 2 and 3, and plays as `tit_for_tat` in other cases

The main problem in CIPD study is not only to find *good* strategies, but also to understand the dynamic of populations of agents using fixed strategies.

2 Ecological tournaments and other simulations

Two kinds of experimentation are used in litterature to evaluate strategies for the CIPD:

- The basic one, is to make a two-by-two round robin tournament between strategies. The payoff of each one would be the total sum of each iterated game[1]. A ranking could then be computed according to the score of each strategy.
  The higher a strategy is ranked, the better it is.
  As shown in previous work, [4], some cycles between strategies may be found (`A` better than `B`, which is better than `C` which is better than `A`), the order created by this method cannot be considered as total.
- The second kind of experimentation is a kind of imitation of the natural selection process, and is closely related to population dynamics, but in a completely discrete context. Let us consider a population of N players, each one adopting a particular strategy. At the beginning we consider that each strategy is equally represented in the population. Then a tournament is made, and good strategies are favoured, whereas bad ones are disadvantaged, by a proportional population redistribution. This redistribution process, also called a generation, is repeated until an eventual population stabilisation, i.e. no changes between two generations.
  A good strategy is then a strategy which stays alive in the population for the longest possible time, and in the biggest possible proportion. This kind of evaluation quotes the *robustness* of strategies.
  This looks like prey-predator model, but is not. The number of involved species is not limited to two, interactions between, or into, species are much more complex, and global population is fixed. Once a population has disappeared it has no way to reappear: there is no stochastic perturbations nor in population distribution, nor in strategies description.

Let us detail the computation method for ecological evolution involving 3 strategies. This will be used to compute results detailed later in this paper.
Suppose that, initially, the population is composed of 3 strategies `A`, `B`, `C`.
At generation n each strategy is represented by a certain number of individuals: $\mathrm{W}_n(\mathtt{A})$ using `A`, $\mathrm{W}_n(\mathtt{B})$ using `B` and $\mathrm{W}_n(\mathtt{C})$ using `C`.
The payoff matrix of two-by-two meeting between `A`, `B` and `C` is computed and is thus known. $\mathrm{V}(\mathtt{A}|\mathtt{B})$ is the score of `A` when it meets `B`, etc.
Let us suppose that the total size of the population is fixed and constant. Let us note it Π.

$$\forall i \in [1, \infty[, \Pi = W_i(\mathtt{A}) + W_i(\mathtt{B}) + W_i(\mathtt{C})$$

The computation of the score (distributed points) of a player using a fixed strategy, at generation n is then :

---

[1] In our experiments every meeting has the same length (1000 moves), but strategies can't guess it.

$$
\begin{aligned}
\mathrm{g}_n(\mathtt{A}) &= W_n(\mathtt{A})V(\mathtt{A}|\mathtt{A}) + W_n(\mathtt{B})V(\mathtt{A}|\mathtt{B}) + W_n(\mathtt{C})V(\mathtt{A}|\mathtt{C}) - V(\mathtt{A}|\mathtt{A}) \\
\mathrm{g}_n(\mathtt{B}) &= W_n(\mathtt{A})V(\mathtt{B}|\mathtt{A}) + W_n(\mathtt{B})V(\mathtt{B}|\mathtt{B}) + W_n(\mathtt{C})V(\mathtt{B}|\mathtt{C}) - V(\mathtt{B}|\mathtt{B}) \\
\mathrm{g}_n(\mathtt{C}) &= W_n(\mathtt{A})V(\mathtt{C}|\mathtt{A}) + W_n(\mathtt{B})V(\mathtt{C}|\mathtt{B}) + W_n(\mathtt{C})V(\mathtt{C}|\mathtt{C}) - V(\mathtt{C}|\mathtt{C})
\end{aligned}
$$

Let us quote that because of the substractions, computation of g cannot be simplified. The total points distributed to all involved strategies are :

$$
\mathrm{t}(n) = W_n(\mathtt{A})\mathrm{g}_n(\mathtt{A}) + W_n(\mathtt{B})\mathrm{g}_n(\mathtt{B}) + W_n(\mathtt{C})\mathrm{g}_n(\mathtt{C})
$$

The size of each sub-population at generation $n+1$ is finally:

$$
W_{n+1}(\mathtt{A}) = \frac{\Pi W_n(\mathtt{A})\mathrm{g}_n(\mathtt{A})}{\mathrm{t}(n)} \tag{3}
$$

$$
W_{n+1}(\mathtt{B}) = \frac{\Pi W_n(\mathtt{B})\mathrm{g}_n(\mathtt{B})}{\mathrm{t}(n)} \tag{4}
$$

$$
W_{n+1}(\mathtt{C}) = \frac{\Pi W_n(\mathtt{C})\mathrm{g}_n(\mathtt{C})}{\mathrm{t}(n)} \tag{5}
$$

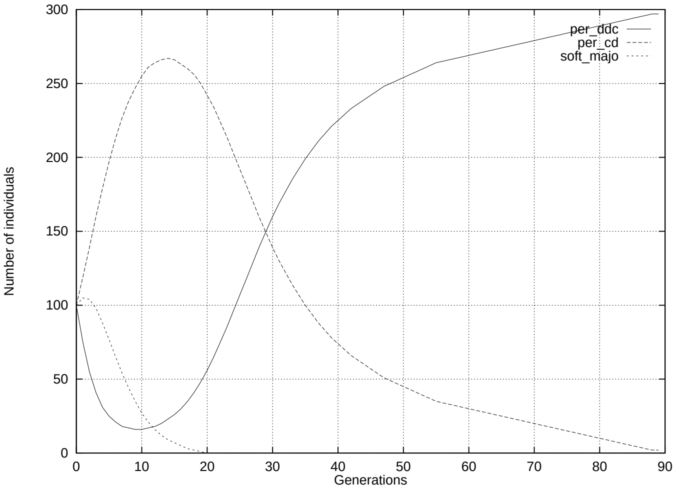
All divisions being rounded to the nearest lower integer.
Classical results on the problem, which have been emphasized by AXELROD in [1], show that to be good a strategy has to:

- be nice, i.e. not be the first to defect[2]
- be reactive
- forgive
- not be too clever, i.e. to be simple in order to be understood by its opponent

The well-known `tit_for_tat` strategy which satisfies all those criteria, has, since AXELROD's book, been considered to be one of the *best* strategies not only for cooperation but also for evolution of cooperation.
It is widely accepted that cooperation seems to be the more general adopted behavior with this model. However it is also clear that there is a gap between those classical results and what appears in the life-as-it-is. There are not only *nice* people. Cooperation, as choice of collective fitness against individual one, is not the emergent stable behavior in all ecosystems.
The model carries those contradictions in some way. For instance, it is possible for a *naughty* strategy population to evoluate better than a nice one, and thus to win an ecological simulation, as seen in the example set up in Fig. 1.
Cooperation has been thought as a *global convergency point* of (artificial) living systems. Definition of ecological evolution, i.e. without any mutation of individuals ; simplicity, and small size, of studied population set ; limitation of computing power, may be explanations of the differences found between formal results and practical constatations.

---

[2] whereas *naughty* strategies defects spontaneously at least one time



	per_cd	per_ddc	soft_majo	Total	Init. Pop.
per_cd	2000	1665	4000	7665	100
per_ddc	2500	1666	671	4837	100
soft_majo	1500	2331	3000	6831	100

Fig. 1. Defectors may be strong

We think that the *simplicity* criteria in the strategy definition is not good, [5], and we have thus introduced a strategy called `gradual`[3], which illustrates our point of view, [2]. We have evaluated it in large environments, [3].

The strategy's point of view is however not the only one to be taken into account when trying to understand evolution of cooperation. Population dynamics play a major role too. So we think it is important to understand it well. In particular it is important to know if *chaotic* dynamics are possible.

It may seem easy to find a particular round robin payoff matrix which could imply complex dynamics in ecological evolutions. But verifying such a matrix corresponds to some strategy sets, and defining those strategies, is harder. The only useful way to study dynamics is then to make systematic exploration.

Previous works, [10, 9, 8], have shown that with stochastic strategies, or under evolutionary conditions, oscillations in evolution of population could often be found. In those particular cases, the stochastic, i.e. non deterministic, element could be one of the main explanations to the oscillatory dynamics observed. Other analytic results are widely spread in the case of population dynamics computed in a not discrete way, see for instance [7].

As we think that to understand complex cases one has to first understand well the behavior of simple, we studied the simplest ones. Thus all strategies we will use in this paper, unlike in [10], are *pure* in the Game Theory meaning. That is they are deterministic.

---

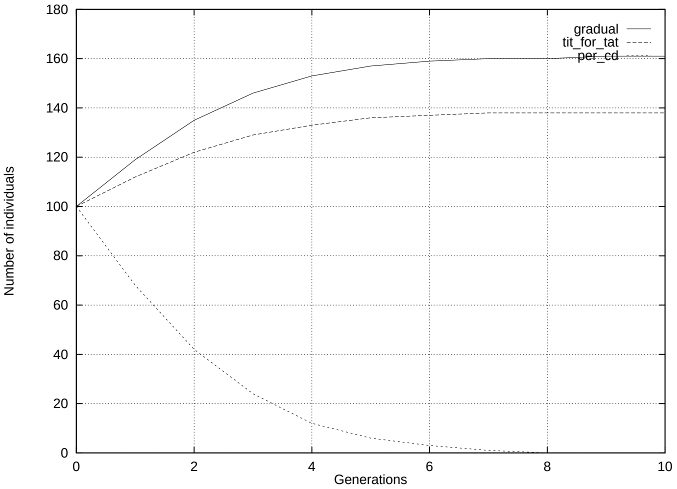
[3] `gradual` cooperates on the first move, then after the first opponent's defection defects one time, and cooperates two times, after the second opponent's defection defects two times and cooperates two time, ..., after the n^{th} opponent's defection defects n times and cooperates two times

3 Unexpected dynamics

In most cases, ecological evolutions look like monotonous convergence, which means that population's evolution curves are always increasing or decreasing. The ranking seems to be determined after few steps. In some cases, however, complex oscillations can be observed. For instance, one can obtain oscillatory movements, which could be attenuated, increasing or periodic. We have undertaken to systematically seek such cases, and we have found some which seems to be at some kind of "*edge of chaos*". It is easier to find oscillations with many strategies but we will see that such oscillations can be found with few ones. In this paper we have chosen to present only situations with three pure strategies involved. After having analysed thousands of evolutions, we propose to classify the phenomena observed in five groups.

3.1 Monotonous convergence

The first of the five groups corresponding to a great majority of cases (99% in our experiments) and which is often thought to be the only one, is that of monotonous convergence. Population's size after such evolution (no or little change in the movement) stabilizes itself completely (see Fig. 2).

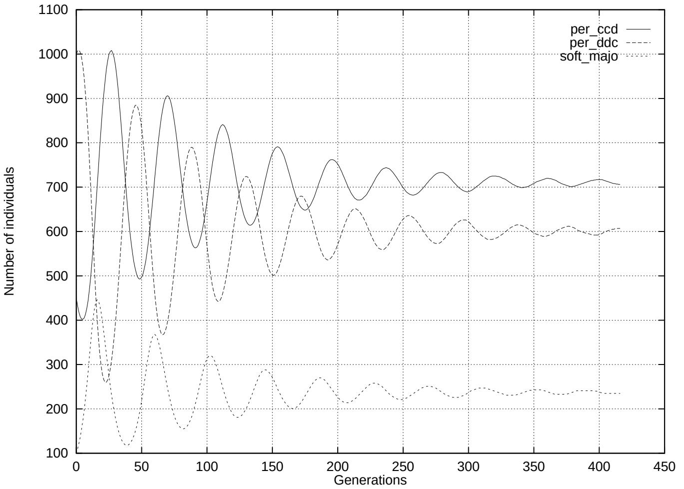


	tit_for_tat	gradual	per_cd	Total	Init. Pop.
tit_for_tat	3000	3000	2498	8498	100
gradual	3000	3000	2955	8955	100
per_cd	2503	615	2000	5115	100

Fig. 2. Monotonous convergence

3.2 Attenuated oscillatory movements

The second case is the attenuated oscillatory movement one. The size of the population oscillates with a decreasing amplitude, which leads at the end of the evolution, as in the first case, to a population stabilization, but this time after many reversals.



	`per_ccd`	`per_ddc`	`soft_majo`	Total	Init. Pop.
`per_ccd`	2334	1665	3666	7665	450
`per_ddc`	3335	1666	671	5672	1000
`soft_majo`	2001	2331	3000	7332	100

Fig. 3. Attenuated oscillatory movements

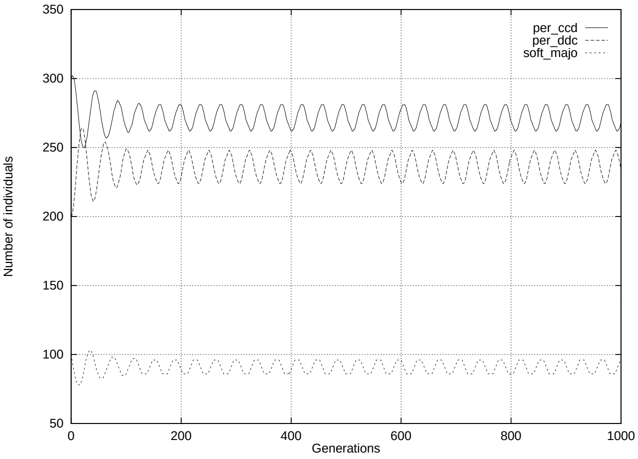
Fig. 3 illustrates this case. Three populations of strategies `per_ccd`, `per_ddc` and `soft_majo` are conflicting themselves with many oscillations during the first 100 generations and then, gradually, find an equilibrium which is reached with generation 420 from which no more modification occurs.

3.3 Periodic movements

The third case is the periodic movements one. Population size of the strategies after a potential phase of hesitation recurringly evolves reproducing after several generations the same combination, without stabilizing (see Fig. 4). On this example, population size comes at the same point every 37 generation. The oscillation is never stabilized.

It seems that such phenomena involves strategies which made a cycle in tournament : `A` is better than `B`, which is better than `C` which is better than `A`. The nature of those relation may be an explanation of those periodic movements.

Such a phenomenon was recently quoted in [12] in the living world and relates to populations of lizards. Even if it is not sure that those results could be applied in the CIPD model, it is useful to notice the coincidence between our three strategies population periodic movements on the one hand, and the real world of terrestrial life on the other hand.



	Init. Pop.
per_ccd	300
per_ddc	200
soft_majo	100

Fig. 4. Periodic movements

3.4 Increasing oscillations

The fourth case, is the increasing oscillations with rupture. The case represented by Fig. 5 is similar to the precedent one, except that now oscillations are growing. It leads at the end to a *break*.

The break is done in profit of `per_ddc` which remains alone, after many oscillations. This kind of dynamics shows that violent oscillations can allow the survival of non-cooperative strategies which benefit of the general disorder.

	Init. Pop.
per_ccd	300
per_ddc	400
soft_majo	200

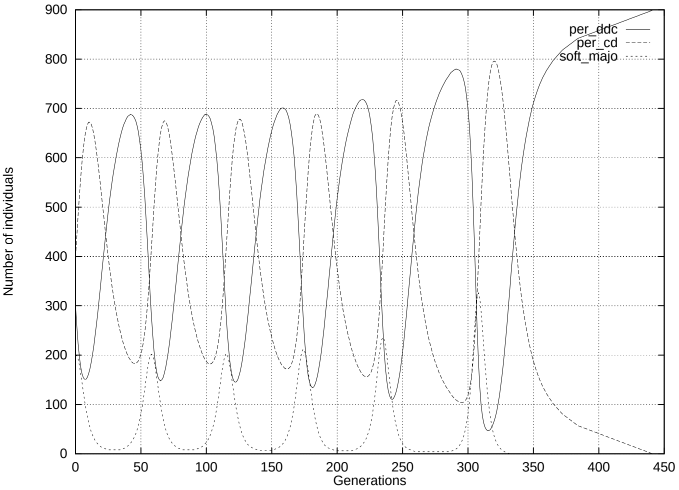
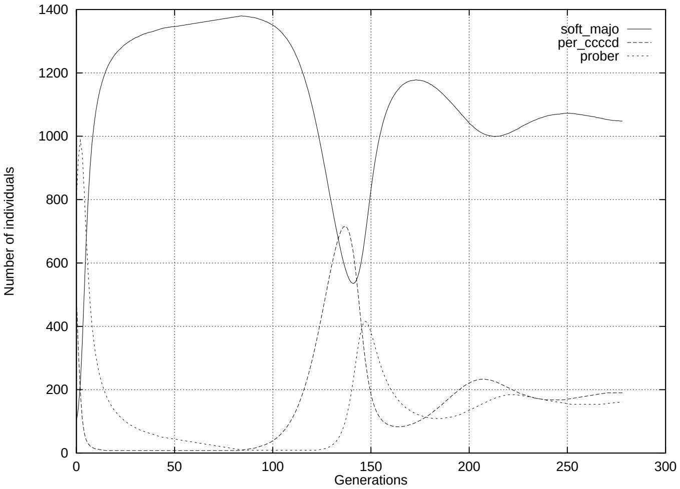


Fig. 5. Increasing oscillations

3.5 Disordered oscillations

The fifth kind of dynamics gathers the cases which don't get into the fourth previous ones. Movements seem disordered. In our experiments these disordered movements do not last long, therefore we hesitate to use the *chaos* term. On Fig. 6, after a strong instability during 250 generations where each of the three strategies comes very close to death, an equilibrium point is reached.



	soft_majo	per_ccccd	prober	Total	Init. Pop.
soft_majo	3000	2400	2999	8399	100
per_ccccd	3400	2600	206	6206	500
prober	2999	4196	1004	8199	800

Fig. 6. Disordered oscillations

4 Sensitivity to initial conditions

In order to try to complete our opinion on the *chaotic aspect* of those dynamics we studied its sensitivity to initial conditions. We found that very small variations of initial parameters could imply important changes in the phenomenom observed.

4.1 Sensitivity to population's size

The transition from a periodic movement to a monotonous one can be made when the initial size of the population varies from one unit.

In the first experiment of Fig. 7, the CIPD parameters are the classical ones (the one represented on Tab. 1, T=5, R=3, P=1 and S=0), each match is 1000 meeting long, there are 300 agents using `per_ccd`, 100 using `soft_majo` and 244 using `per_ddc`. Populations evoluate in a periodic movement.

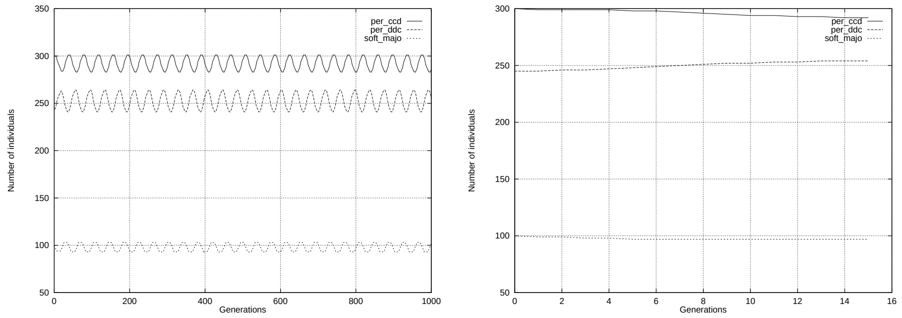


Fig. 7. Sensitivity of dynamics to population's size. All parameters are identical except for the initial size of `per_ddc` which is 244 on the left and 245 on the right

If only one `per_ddc` agent, which is a *naughty* one, is added then the evolution is a monotonous convergence.
Variation of one unit in the initial population of a strategy can also change the winner of the ecological evolution.
In the experiment of the Fig. 8, the conditions are the same as for the previous ones, except for the size of population. There are 100 `per_ddc`, 159 `soft_majo`, and 100 `per_cd`. The winner is `per_ddc`. If only one `soft_majo` is added, then the winner is `per_cd`. It could be noticed that the modified strategy **never** wins in any cases.

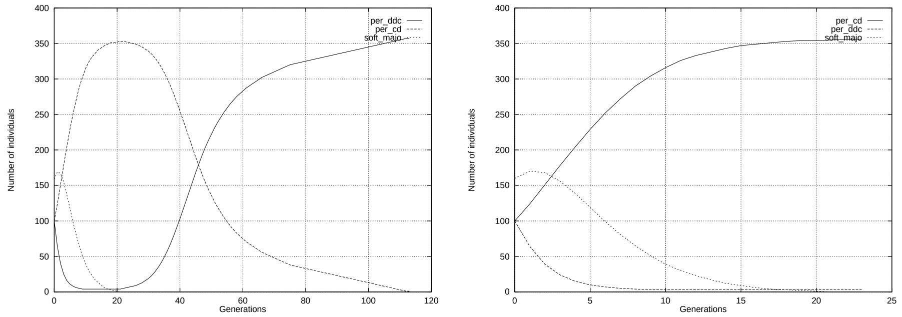


Fig. 8. Sensitivity of winner to population's size. All parameters are identical except for the initial size of `soft_majo` which is 159 on the left and 160 on the right.

4.2 Sensitivity to game length

A change of the kind of dynamic can be created by the variation of the length of game (number of iteration of the Prisoner's Dilemma), which is fixed but unknown by strategies.
In the experiment of Fig. 9 the CIPD is used (5,3,1,0). There are 300 `per_ccd`, 100 `soft_majo`, and 244 `per_ddc`. When the game lasts 7 moves the dynamic is

a periodic movement and becomes an attenuated oscillary one when the game is 6 moves long.

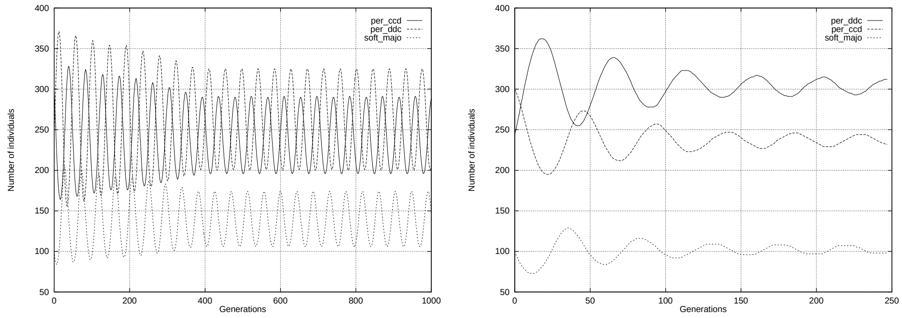


Fig. 9. Sensitivity to game length. All parameters are identical except for the game length which is 7 moves on the left and 6 on the right.

4.3 Sensitivity to CIPD payoff

A change in the Prisoner's Dilemma payoff matrix, with respect to inequations 1 and 2, may change the kind of dynamic.

In the experiment of Fig. 10 there are 300 `per_ccd`, 100 `soft_majo` and 244 `per_ddc`. Games last 1000 moves. R, P, and S are the same as in the classical choice of the Tab. 1, but T=4.6. The dynamic is an increasing oscillation movement. When T=4.7 then it comes to a periodic one.

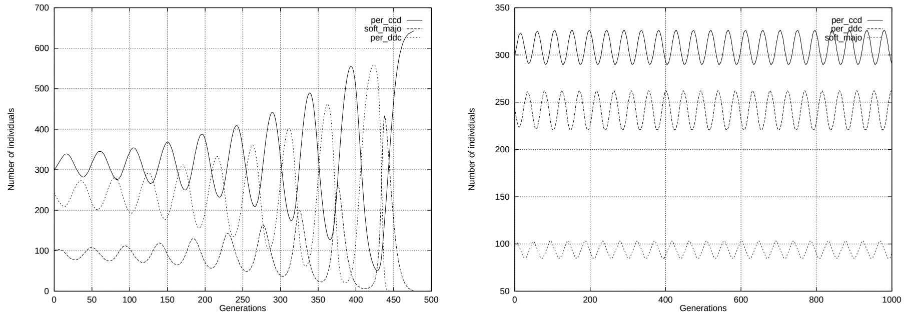


Fig. 10. Sensitivity to CIPD payoff. All parameters are identical except that T=4.6 on the left and T=4.7 on the right.

4.4 Sensitivity to repartition computation method

Dynamics can change with the individual repartition method used in ecological evolution between two generations (rounding computation method).

In the experiment of Fig. 11 the CIPD parameters are the classical one (5,3,1,0), games are 1000 moves long, there are 300 `per_ccd`, 100 `soft_majo` and 200 `per_ddc`. If we round the number of individuals as in equations 3, 4 and 5 the dynamic is a periodic movement, whereas if we use real values (populations are no more discrete), it comes to an attenuated oscillation one.

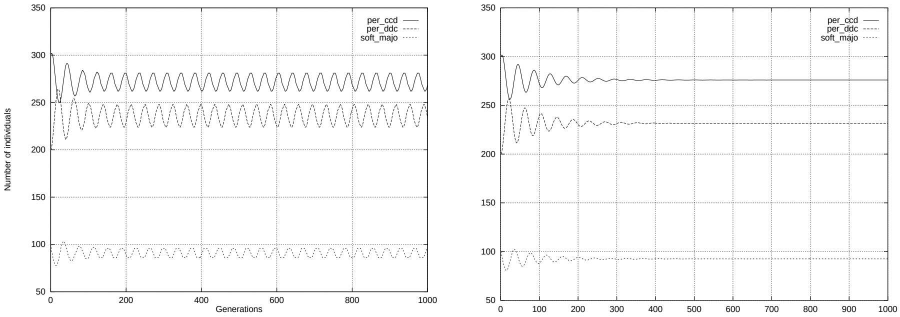


Fig. 11. Sensitivity to repartition computation method. All parameters are identical except that repartition on the left is done by rounding and uses real value on the right.

Another proof of this sensitivity to rounding computation method is that when multiplying all proportions by a constant factor, dynamic changes.
In the experiment of Fig. 12, the CIPD parameters are the classical ones (5,3,1,0), games are 1000 moves long, there are 450 `per_ccd`, 100 `soft_majo` and 1000 `per_ddc`. Dynamic is an attenuated oscillating movement. If all populations are divided by 10, it becomes an increasing one.

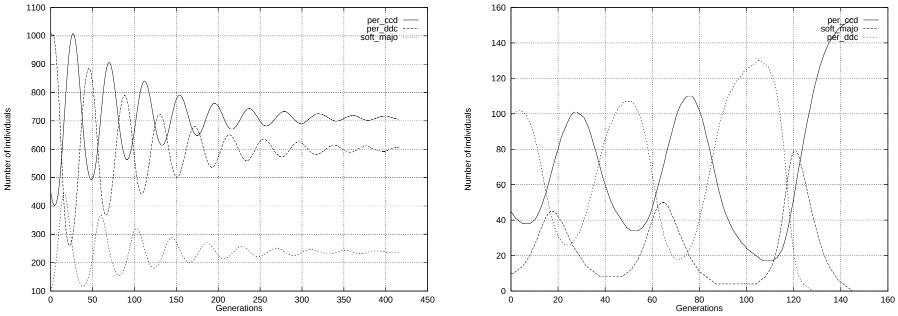


Fig. 12. Sensitivity to repartition computation method. All parameters are identical except that populations on the right are divided by 10.

We also note that small modifications in the composition of strategies sometimes involve the disappearance of the oscillation, or change the winner of an increasing oscillation. In most of the cases the shape of the curves is different.

5 Conclusion

In some very rare and particular cases, ecological evolution comes to disordered population dynamics, which we classify into 5 categories. In the periodic cases, with attenuated or increasing oscillations, there are always some *naughty* strategies involved, as expected. Sometimes, during the break of an attenuated oscillation the winner is a *naughty* one. It seems that disorder gives more chance to *not nice* strategies and is unfavourable to cooperative ones.
Instability to initial conditions, and the fact that the winner is not always a *nice* one, makes the way the population evolves almost unpredictable if using only round robin tournament results and ecological evolution equations.
It is possible that with highly complex strategies, what seems exceptionnal here becomes more frequent. If it is the case, this would mean that, unlike the most accepted interpretation of CIPD, cooperation is not the most frequent attractor state, when agents with complex behavior are involved.
Social relations would then be unstable by nature because of oscillatory dynamics leading to break, which benefits to aggressive strategies.
Simulation software with many strategies is already available for all plateforms, through our web site at `http://www.lifl.fr/IPD` or by anonymous ftp at `ftp.lifl.fr` in `pub/projects/IPD`.

References

[1] R. Axelrod. *The Evolution of Cooperation.* Basic Books, New York, USA, 1984.

[2] B. Beaufils, J.P. Delahaye, and P. Mathieu. Our Meeting with Gradual, A Good Strategy for the Iterated Prisoner's Dilemma. In Christopher G. Langton and Katsunori Shimohara, editors, *Proceedings of the Fifth International Workshop on the Synthesis and Simulation of Living Systems*, pages 202–209, Cambridge, MA, USA, 1996. The MIT Press/Bradford Books. Artificial Life 5, Nara, Japan, May 16-18 1996.

[3] Bruno Beaufils, Jean-Paul Delahaye, and Philippe Mathieu. Complete classes of strategies for the classical iterated prisoner's dilemma. In V. W. Porto, N. Saravanan, D. Waagen, and A. E. Eiben, editors, *Evolutionnary Programming VII*, volume 1447 of *Lecture Notes in Computer Science*, pages 33–41, Berlin, 1998. Springer-Verlag. Evolutionnary Programing VII, San Diego, CA, USA, March 25-27, 1998.

[4] J.P. Delahaye and P. Mathieu. Expériences sur le dilemme itéré des prisonniers. Publication Interne IT-233, Laboratoire d'Informatique Fondamentale de Lille, Lille, France, 1992.

[5] J.P. Delahaye and P. Mathieu. Complex Strategies in the Iterated Prisoner's Dilemma. In A. Albert, editor, *Chaos and Society*, volume 29 of *Frontiers in Artificial Intelligence and Applications*, pages 283–292, Amsterdam, Netherlands, 1995. Université du Québec à Hull, Canada, IOS Press/Presses de l'Université du Québec. Chaos & Society 1994, Trois Rivières, Canada, June 1-2 1994.

[6] Merill M. Flood. Some experimental games. Research memorandum RM-789-1-PR, RAND Corporation, Santa-Monica, CA, USA, June 1952.

[7] Josef Hofbauer and Karl Sigmund. *Evolutionary Games and Population Dynamics.* Cambridge University Press, Cambridge, UK, 1998.

[8] K. Lindgren. Evolutionary Phenomena in Simple Dynamics. In Christopher G. Langton, Charles Taylor, J. Doyne Farmer, and Steen Rasmussen, editors, *Artificial Life II: Proceedings of the Second Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems*, volume 10 of *Santa Fe Institute Studies in the Sciences of Complexity*, pages 295–312, Reading, MA, USA, 1992. Addisson–Wesley Publishing Company. Artificial Life 2, Santa Fe, USA, February 1990.

[9] John H. Nachbar. Evolution in the finitely repeated prisoner's dilemma. *Journal of Economic Behavior and Organization*, 19:307–326, 1992.

[10] M. Nowak and K. Sigmund. Oscillations in the Evolution of Reciprocity. *Journal of Theoretical Biology*, 137:21–26, 1989.

[11] W. Poundstone. *Prisoner's Dilemma : John von Neumann, Game Theory, and the Puzzle of the Bomb*. Oxford University Press, Oxford, UK, 1993.

[12] J. Maynard Smith. The Games the Lizards Play. *Nature*, 380:198–199, March 1996.

[13] John von Neumann and Oskar Morgenstern. *Theory of Games and Economics Behavior*. Princeton University Press, Princeton, NJ, USA, 1944.