New Winning Strategies for the Iterated Prisoner's Dilemma

Philippe Mathieu[1] **and Jean-Paul Delahaye**[1]

[1] *CRISTAL Lab 9189 - CNRS, Lille University, Villeneuve d'Ascq, France*
Correspondence should be addressed to *philippe.mathieu@univ-lille.fr*

Journal of Artificial Societies and Social Simulation 20(4) 12, 2017
Doi: 10.18564/jasss.3517 Url: http://jasss.soc.surrey.ac.uk/20/4/12.html

Received: 08-09-2016 Accepted: 02-08-2017 Published: 31-10-2017

Abstract: In the iterated prisoner's dilemma game, new successful strategies are regularly proposed especially outperforming the well-known *tit_for_tat* strategy. New forms of reasoning have also recently been introduced to analyse the game. They lead William Press and Freeman Dyson to a double infinite family of strategies that -theoretically- should all be efficient strategies. In this paper, we study and confront using several experimentations the main strategies introduced since the discovery of *tit_for_tat*. We make them play against each other in varied and neutral environments. We use the `complete classes` method that leads us to the formulation of four new simple strategies with surprising results. We present massive experiments using simulators specially developed that allow us to confront up to 6,000 strategies simultaneously, which had never been done before. Our results show without any doubt the most robust strategies among those so far identified. This work defines new systematic, reproductible and objective experiments suggesting several ways to design strategies that go a step further, and a step in the software design technology to highlight efficient strategies in iterated prisoner's dilemma and multiagent systems in general.

Keywords: Game Theory, Group Strategy, Iterated Prisoner's Dilemma (IPD), Agent's Behaviour, Memory, Opponent Identification

Introduction

1.1 The iterated prisoner's dilemma is a game that allows to understand various basic truths about social behaviour and how cooperation between entities is established and evolves sharing same space: living organisms sharing an ecological niche, companies competitors fighting over a market, people with questions about the value of conducting a joint work, etc (Axelrod 2006; Beaufils & Mathieu 2006; Kendall et al. 2007; Mathieu et al. 1999; Mittal & Deb 2009; Poundstone 1992; Rapoport & Chammah 1965; Sigmund 2010). Although based on an extreme simplification of the interactions between entities, the mathematical study of the iterated prisoner's dilemma remains difficult, and often, only computer simulations are able to solve classical questions or identify ways of building efficient behaviours (Beaufils et al. 1996; Kendall et al. 2007; Li et al. 2011; Mathieu et al. 1999; Tzafestas 2000).

1.2 A series of works (Beaufils et al. 1996; Delahaye et al. 2000; LI & Kendall 2013; Li et al. 2011; O'Riordan 2000; Press & Dyson 2012; Tzafestas 2000; Wedekind & Milinski 1996) have introduced other efficient strategies than the famous *tit_for_tat*. Each time, discovered strategies have been justified by mathematical or experimental arguments trying to establish that we are dealing with better strategies than *tit_for_tat*. These arguments are often convincing, but however, they do not help to highlight a strategy that can be unanimously considered better than the others. It is not even possible today to know what are among the best fifteen strategies identified, those actually in the top, and what are the right elements for structuring efficient and robust behaviour. We have begun to study the actual situation with the desire to reach clear and as unbiased as possible conclusions.

1.3 Our method is based on three main ideas, each converging toward robust results and objectives aims.

1. Confronting the candidate strategies on a `tournament` (mainly for information) and the method of `evolutionary competition` which leads to results partially independent from initial conditions.

2. Using sets of strategies coming from a particular class (eg using the last move of past of each player) are in competition. This method of `complete classes` (Beaufils et al. 1998) avoids the subjective choice usually done when one tries to build his own set of strategies. We use these classes in two ways. First, we use them alone (thus without any added strategy) , thereby objectively identify efficient strategies, and secondly we complement with sets of the most successful strategies that we want to compete and rank. This allows us to identify the most robust and resilient strategies.

3. Taking an incremental approach , combining the results of several progressive series of massive confrontation experiments in order to be able to formulate, as closely as possible, robust conclusions.

1.4 Our aim in this paper is to identify new systematic, reproducible and objective experiments, suggesting several ways to design robust and efficient new strategies and more than that, a general scheme to identify new ones.

1.5 This experimental method has no known theoretical equivalent. Indeed, for iterated games in general, but especially for the iterated prisoner's dilemma, notions of Nash equilibrium, Pareto optimality or evolutionarily stable strategies Lorberbaum (1994); Lorberbaum et al. (2002) do not suggest new and efficient strategies and have never led to discover any new interesting strategy. One will find in Wellman (2006) other paths to follow that would lead to strengthen our results or add new ones. This field is quite difficult to study theoretically.

1.6 One of the obvious reasons is that it is impossible to make the optimal score against all strategies. This is a consequence of the first move: to play optimally against *all_d* it is necessary to defect at the first round, to play optimally against *spiteful* it is necessary to cooperate. Another reason comes from the infinite set of possible strategies, not endowed with a natural topology. The approach by evolutionary algorithms do not seem to work and never reveal any new robust strategy. The incremental method described in this paper allows to discover new behaviours and unexpected simple strategies.

1.7 In Section 2 we recall the rules of the iterated prisoner's dilemma and specially tournaments and evolutionary competitions used to evaluate strategies. In Section 3 we define precisely well known classical deterministic strategies and several probabilistic ones coming from the state of the art, and evaluate them both in tournaments and evolutionary competitions. In Section 4 we present the complete classes principal which is an objective frame to find and compare strategies : the main idea is to build a set of all the possible strategies using the same size of memory. In Section 5 we show all the results we can identify with these complete classes alone. Using these results we identify four promising new strategies. In Section 6 we confront all the strategies defined during the previous sections all together mainly to test robustness of the best ones.

1.8 This paper is a completed and extended version of the two page paper Mathieu & Delahaye (2015). All the strategies, experiments and mainly the whole package allowing to replicate reported simulation experiments can be downloaded on our web site `http://www.lifl.fr/IPD/ipd.html`.

Rules of the Game

2.1 The prisoner's dilemma is that accorded to two entities with a choice between cooperation (`c`) and defection (`d`) and are remunerated by `R` points each if each plays `c`, `P` points if each plays `d` and receiving `T` respectively `S` points if one plays `d` and the other `c`. We describe these rules by writing:
`[c, c] -> R + R`, `[d, d] -> P + P`, `[d, c] -> T + S`.

2.2 We require that $T > R > P > S$ and $T + S < 2R$ The classical chosen values are `T = 5, R = 3, P = 1, S = 0`, which gives: `[c, c] -> 3 + 3`, `[d, d] -> 1 + 1`, `[d, c] -> 5 + 0`.

		Player II	
		Cooperate	Defect
Player I	Cooperate	R=3 R=3	T=5 S=0
	Defect	S=0 T=5	P=1 P=1

2.3 It is a dilemma situation because both entities can collectively win 6 points playing `[c, c]`, whereas they win less playing `[c, d]` and even less playing `[d, d]`. The collective interest is that everyone play `c`, but a single logical analysis leads inevitably to `[d, d]` which is collectively the worst case !

2.4 The dilemma is iterated when we imagine that the situation of choice between `c` and `d` is presented periodically to the same two entities. The game consists in choosing a strategy that, informed about the past (hence the previous behaviour of the opponent), shows how to play the next move. We recall that in this game, we cannot play well against everyone. Playing well against *all_d* need to always betray (and in particular for the first move), and playing well against *all_c* need to always cooperate. But moves being simultaneous, one cannot play optimally against these two strategies.

2.5 In this game, since winning against everyone is trivial (*all_d* does), it is obvious that "playing well" corresponds to earning a maximum of points, which in evolutionary competitions is equivalent to ending with the greatest population possible.

2.6 When a set *A* of strategies is given, we can evaluate it in two ways to get a ranking.

- Tournaments: each strategy meets each other (including itself) during a series of *n* moves (we take *n = 1,000* in the experiments below). Accumulated points earned by a strategy give its score (which thus depends on *A*). The ranking is in respect with the scores.
- Evolutionary competition(Axelrod 2006): take a number *K* of strategies of each kind in *A* (eg *K = 100*), which is what is known as Generation 1, `G1`. A tournament between the strategies `G1` determines the scores of each strategy. Each strategy will have in generation 2, `G2`, a number of descendants proportional to its score and only those descendants constitute generation 2, `G2`. It is assumed that the total number of strategies remains constant from one generation to the next (Cardinal (G1) = Cardinal (G2) = ...). Generation 3, `G3`, is calculated from the same 2nd generation etc. In an evolutionary competition, strategies that are playing poorly are quickly eliminated. Therefore, those exploiting some strategies playing poorly (which can be numerous especially in complete classes) soon stop to take benefit of them. Finally one can note that only survive the strategies playing efficiently against strategies playing efficiently too.

The Basic Strategies

3.1 We make a distinction between deterministic strategies and probabilistic strategies, where choices can depend on chance.

3.2 The study of literature about the dilemma led us to define a set of 17 basic deterministic strategies (including the simplest imaginable strategies). We have added 13 probabilistic strategies mainly taking into account the recent discoveries of Press and Dyson on extortion (Press & Dyson 2012).

3.3 Let us present the set of 17 basic strategies

all_c always cooperates

all_d always defects

tit_for_tat cooperates on the first move then plays what its opponent played the previous move (Rapoport & Chammah 1965).

spiteful cooperates until the opponent defects and thereafter always defects (Axelrod 2006). Sometimes also called *grim*.

soft_majo begins by cooperating and cooperates as long as the number of times the opponent has cooperated is greater that or equal to the number of times it has defected. Otherwise she defects (Axelrod 2006).

hard_majo defects on the first move and defects if the number of defections of the opponent is greater than or equal to the number of times she has cooperated. Else she cooperates (Axelrod 2006).

per_ddc plays ddc periodically

per_ccd plays ccd periodically

mistrust defects on the first move then play what my opponent played the previous move (Axelrod 2006). Sometimes also called *suspicious_tft*.

per_cd plays cd periodically

pavlov cooperates on the first move and defects only if both the players did not agree on the previous move (Wedekind & Milinski 1996). Also called *win-stay-lose-shift*.

tf2t cooperates the two first moves, then defects only if the opponent has defected during the two previous moves (Some authors call it sometimes erroneously hard_tft. These is often a confusion between these two).

hard_tft cooperates the two first moves, then defects only if the opponent has defected one of the two previous moves

slow_tft cooperates the two first moves, then begin to defect after two consecutive defections of its opponent. Returns to cooperation after two consecutive cooperations of its opponent.

gradual Cooperates on the first move, then defect n times after n^{th} defections of its opponent, and calms down with 2 cooperations (Beaufils et al. 1996).

prober plays the sequence `d,c,c`, then always defects if its opponent has cooperated in the moves 2 and 3. Plays as *tit_for_tat* in other cases (Mathieu et al. 1999).

mem2 behaves like *tit_for_tat*: in the first two moves, and then shifts among three strategies *all_d*, *tit_for_tat*, *tf2t* according to the interaction with the opponent on last two moves:
A: if the payoff in the two moves is `2R` (`[c,c]` and `[c,c]`) then *tit_for_tat* in the following two moves
B: if the payoff in the last move is `T+S` (`[c,d]` or `[d,c]`) then *tf2t* in the following 2 moves
C: in all other cases she plays *all_d* in the following two moves
D: if *all_d* has been chosen twice, she always plays *all_d*. (LI & Kendall 2013)

3.4 Let us present now a set of 12 probabilistic strategies. These strategies start with `c`, then play `c` with probability
- p_1 if the last move is `[c,c]`
- p_2 if the last move is `[c,d]`
- p_3 if the last move is `[d,c]`
- p_4 if the last move is `[d,d]`

equalizerA	$p_1 = 3/4$	$p_2 = 1/4$	$p_3 = 1/2$	$p_4 = 1/4$
equalizerB	$p_1 = 9/10$	$p_2 = 7/10$	$p_3 = 1/5$	$p_4 = 1/10$
equalizerC	$p_1 = 9/10$	$p_2 = 1/2$	$p_3 = 1/2$	$p_4 = 3/10$
equalizerD	$p_1 = 27/35$	$p_2 = 17/35$	$p_3 = 1/5$	$p_4 = 2/35$
equalizerE	$p_1 = 2/3$	$p_2 = 0$	$p_3 = 2/3$	$p_4 = 1/3$
equalizerF	$p_1 = 1$	$p_2 = 13/15$	$p_3 = 1/5$	$p_4 = 2/5$
extortionA	$p_1 = 8/9$	$p_2 = 2/9$	$p_3 = 11/18$	$p_4 = 0$
extortionB	$p_1 = 4/5$	$p_2 = 1/10$	$p_3 = 3/5$	$p_4 = 0$
extortionC	$p_1 = 11/12$	$p_2 = 5/24$	$p_3 = 2/3$	$p_4 = 0$
extortionD	$p_1 = 5/6$	$p_2 = 1/4$	$p_3 = 1/2$	$p_4 = 0$
extortionE	$p_1 = 17/20$	$p_2 = 3/40$	$p_3 = 7/10$	$p_4 = 0$
extortionF	$p_1 = 11/15$	$p_2 = 2/15$	$p_3 = 7/15$	$p_4 = 0$

3.5 These 12 strategies have been chosen randomly among the infinity of possible choices, for no reason other than to obtain a sample as diverse as possible. *Equalizers* and *Extortions* have been introduced in Press & Dyson (2012) and are among strategies called *Zero-Determinant* (ZD) strategies. A *ZD* strategy can enforce a fixed linear relationship between expected payoff between two players. Extortion strategies ensure that an increase in one's own payoff exceeds the increase in the other player's payoff by a fixed percentage. Extortion is therefore able to dominate any opponent in a one-to-one meeting. Equalizer strategies ensure to the other player any payoff between `P` and `R`.

3.6 We conclude this set with the **random** strategy , playing 50% `c`, 50% `d`

Evaluation of the 17 basic strategies

3.7 The experiment `Exp1`, is done using the 17 basic strategies and leads to the following results:

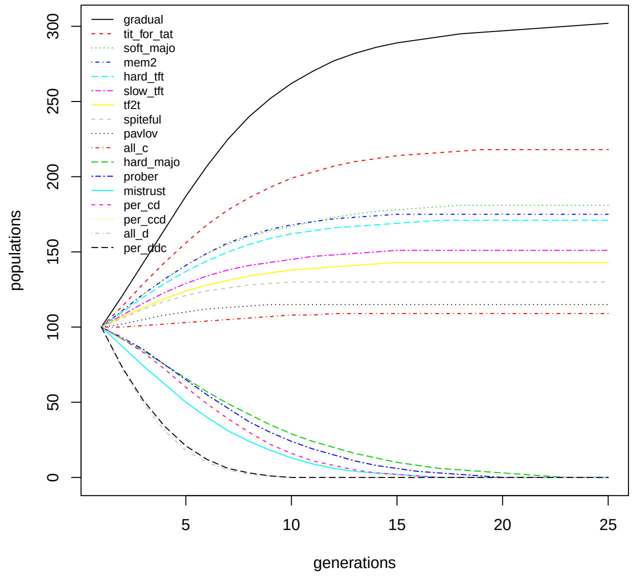

Figure 1: Evolutionary competition Exp1 involving 17 basic strategies. The X-axis represents generations, the Y-axis represents populations. Each point gives the population of the given strategy at the corresponding generation.

Tournament ranking

1	gradual	48827
2	tit_for_tat	46161
3	mem2	45006
4	soft_majo	44830
5	hard_tft	44671
6	slow_tft	43824
7	tf2t	43159
8	spiteful	43003
9	pavlov	41420
10	all_c	40500
11	prober	37688
12	per_ccd	37512
13	per_cd	37392
14	hard_majo	37351
15	mistrust	35197
16	per_ddc	29629
17	all_d	29116

Evolutionary ranking

1	gradual	302
2	tit_for_tat	218
3	soft_majo	181
4	mem2	175
5	hard_tft	171
6	slow_tft	151
7	tf2t	143
8	spiteful	130
9	pavlov	115
10	all_c	109
11	hard_majo	0
12	prober	0
13	mistrust	0
14	per_cd	0
15	per_ccd	0
16	all_d	0
17	per_ddc	0

3.8 Note that in all the evolutionary rankings presented in this papers the order of the strategies is determined by the survival population, and if not, by the time of death.

3.9 The result of the meetings (Tournament and Evolutionary competition) of this set of 17 classic deterministic strategies is a really good validity test of any IPD simulator.

Evaluation of the 30 (17+13) strategies

3.10 Experiment Exp2 uses the 30 deterministic and probabilistic strategies and leads to the following results (Only the 10 first strategies are given):

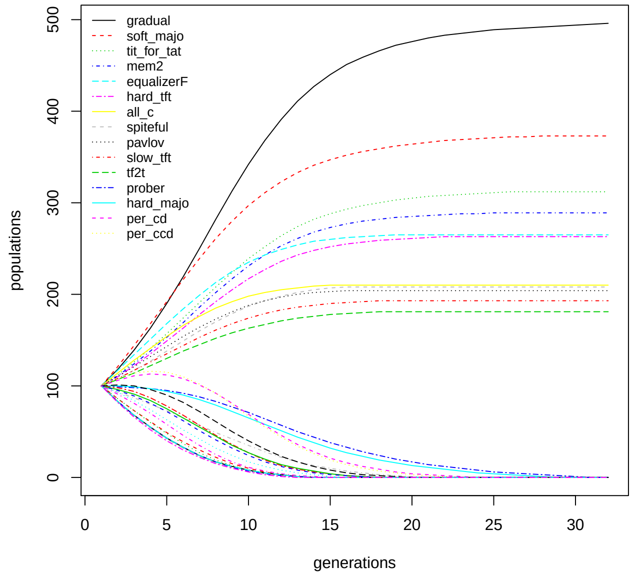

Figure 2: Evolutionary competition Exp2 involving 30 strategies built with 17 deterministic strategies and 13 probabilistic strategies.

Tournament ranking			**Evolutionary ranking**		
1	soft_majo	3653406	1	gradual	496
2	gradual	3562555	2	soft_majo	373
3	equalizerF	3519682	3	tit_for_tat	312
4	all_c	3448902	4	mem2	289
5	tit_for_tat	3373188	5	equalizerF	265
6	mem2	3351097	6	hard_tft	263
7	pavlov	3331900	7	all_c	210
8	hard_tft	3318105	8	spiteful	208
9	slow_tft	3262610	9	pavlov	204
10	per_ccd	3252819	10	slow_tft	193

3.11 We emphasize that for each tournament including a probabilistic strategy, the tournament is always repeated 50 times.

3.12 We note that the only strategy that appears coming from Press and Dyson ideas is the *equalizerF* strategy, that we will encounter often further. It reveals itself the fifth of the 30 strategies here on a competitive basis.

The Complete Classes Principle

4.1 We define the *memory(X,Y)* complete class which is the class of all deterministic strategies using my *X* last moves and the *Y* last moves of my opponent.

4.2 In each *memory(X,Y)* complete class, all deterministic strategies can be completely described by their "genotype" i.e. a chain of C/D choices that begin with the $max(X, Y)$ first moves (not depending on the past). These starting choices are written in lower case. The list of cases of the past is sorted by lexicographic order on my *X* last moves (from the older to the newer) followed by my opponent's *Y* last moves (from the older to the newer) . Here is the genotype of a *memory(1,2)* strategy noted *mem12_ccCDCDDCDD* also called below *winner12*.

My two first plays: c, c

Me-1	She-2	She-1	
C	C	C	C
C	C	D	D
C	D	C	C
C	D	D	D
D	C	C	D
D	C	D	C
D	D	C	D
D	D	D	D

4.3 We indicate the number of strategies we can define in each memory class. Each *memory(X,Y)* class contains a large number of memXY_... strategies. The general formula for the number of elements of a *memory(X,Y)* complete class is $2^{max(X,Y)}.2^{2^{X+Y}}$.

Name	Size
memory(0,1)	$2^1 * 2^2 = 8$
memory(1,0)	$2^1 * 2^2 = 8$
memory(1,1)	$2^1 * 2^4 = 32$
memory(2,0)	$2^2 * 2^4 = 64$
memory(1,2)	$2^2 * 2^8 = 1024$
memory(2,1)	$2^2 * 2^8 = 1024$
memory(2,2)	$2^2 * 2^{16} = 262144$

4.4 Many well known strategies can be defined with this kind of genotype:

all_c	= mem00_C
all_d	= mem00_D
per_cd	= mem10_cDC
per_dc	= mem10_dDC
tit_for_tat	= mem01_cCD
mistrust	= mem01_dCD
spiteful	= mem11_cCDDD
pavlov	= mem11_cCDDC
tf2t	= mem02_ccCCCD
hard_tft	= mem02_ccCDDD
slow_tft	= mem12_ccCCCDCDDD

4.5 Let X, X', Y, Y' be four integers with $X \leq X'$ and $Y \leq Y'$,

If $max(X,Y) = max(X',Y')$ then
$$memory(X,Y) \subseteq memory(X',Y').$$

4.6 Take care that if $max(X,Y) \neq max(X',Y')$ then there is no inclusion because of the beginning.

4.7 That means that if one increases the $min(X,Y)$ of a memory class, not more than the max, then all the *memory(X,Y)* are always in the increased class. For example *memory(0,3)* ⊂ *memory (1,3)* ⊂ *memory (2,3)* ⊂ *memory (3,3)* but not in *memory(0,4)*.

4.8 We can note that several different genotypes can describe finally the same behaviour. For example, the *all_d* strategy appears four times in the memory(1,1) complete class: *mem11_dCCDD* , *mem11_dCDDD*, *mem11_dDCDD*, *mem11_dDDDD*

4.9 Our theoretical hypothesis is that the better you are in a complete class, and the larger the class is, the more chances you have of being robust. Indeed the extent of the complete class guarantees a high degree of behavioral variability without the slightest subjective bias to which one could not escape if one chooses one by one the strategies that one puts in competition.

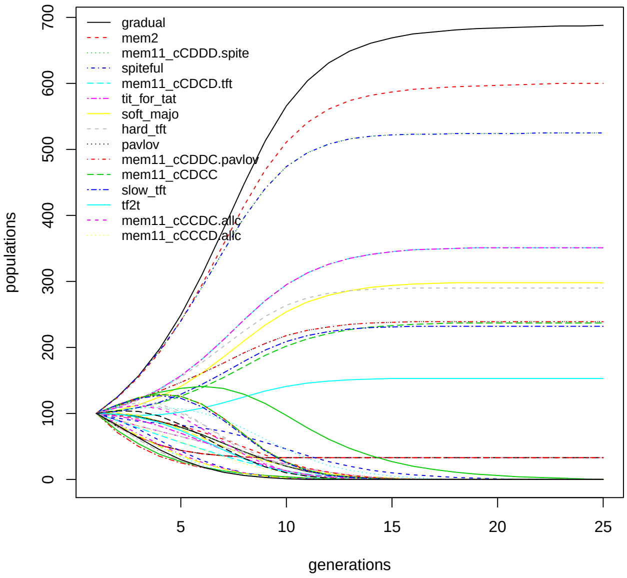

Figure 3: Evolutionary competition Exp3 involving 49 strategies built with 17 basic strategies + 32 *memory(1,1)*.

17 Deterministic and memory(1,1)

4.10 If we consider only deterministic strategies making their decision using the last move of each player, we can define a set of 32 strategies, each determined by a 5-choice genotype $C_1C_2C_3C_4C_5$.
C_1 = move chosen at first when no information is available.
C_2 = move chosen when last move was [c,c]
C_3 = move chosen when last move was [c,d]
C_4 = move chosen when last move was [d,c]
C_5 = move chosen when last move was [d,d]

4.11 Some strategies for this complete class are already among the 30 basic strategies that we have adopted. Some strategies with different genotypes yet still behave identically. We have not sought to remove these duplicates because it makes very small difference to the results, and when we consider larger complete classes it becomes almost impossible.

4.12 Exp3 experiment uses the 17 basic deterministic strategies and the 32 strategies coming from the complete class *memory(1,1)*. This leads to a set of 49 strategies.

	Tournament ranking			Evolutionary ranking	
1	spiteful	138931	1	gradual	688
	mem11_cCDDD-spite	138931	2	mem2	600
3	gradual	138689	3	mem11_cCDDD-spite	525
4	mem2	136928		spiteful	525
5	all_d	125116	5	mem11_cCDCD-tft	351
	mem11_dCCDD-alld	125116		tit_for_tat	351
	mem11_dCDDD-alld	125116	7	soft_majo	298
	mem11_dDCDD-alld	125116	8	hard_tft	290
	mem11_dDDDD-alld	125116	9	pavlov	239
10	mem11_cDDDD	125083		mem11_cCDDC-pavlov	239

4.13 The *all_d* strategy that goes well ranked during the tournament, disappears from the top ten of the evolutionary competition. It's easy to find an explanation: *all_d* exploits strategies playing poorly (nonreactive for example); when they are gone, *all_d* is not able to win enough points to survive.

All Basics and memory(1,1)

4.14 Now in Exp4 we take all the basic strategies (deterministic and probabilistic) with the 32 of the complete class

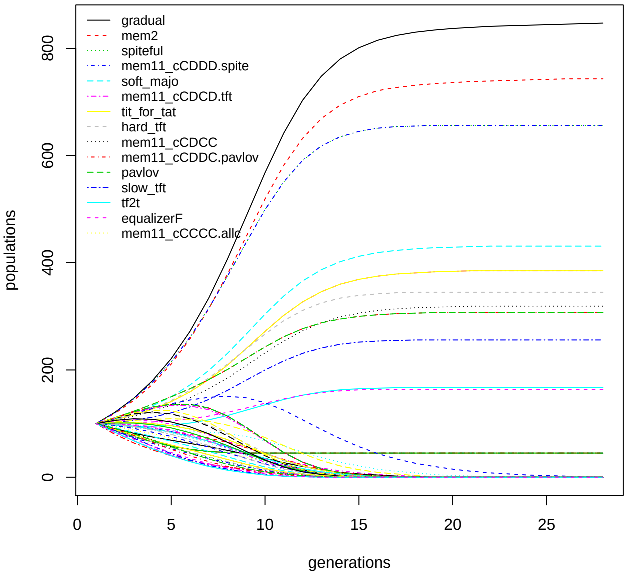

Figure 4: Evolutionary competition Exp4 involving 62 strategies built with 17 basic strategies + 13 probabilistic strategies + 32 *memory(1,1)*.

memory(1,1). This builds a set of 62 (= 17 + 13 + 32) strategies.

	Tournament ranking	
1	gradual	8060369
2	mem11_cCDDD-spite	8050092
3	spiteful	8046164
4	mem2	7946036
5	mem11_dDDDC	7452299
6	pavlov	7425552
7	mem11_cCDDC-pavlov	7422521
8	mem11_dDCDD-alld	7404816
9	mem11_dDDDD-alld	7404016
10	mem11_dCCDD-alld	7403460

	Evolutionary ranking	
1	gradual	850
2	mem2	742
3	mem11_cCDDD-spite	658
4	spiteful	656
5	soft_majo	431
6	tit_for_tat	385
7	mem11_cCDCD-tft	385
8	hard_tft	346
9	mem11_cCDCC	319
10	mem11_cCDDC-pavlov	307

4.15 This ranking confirms that the strategies we have adopted are effectively efficient strategies. The strategies *gradual*, *spiteful* and *mem2* are the three winners: they are good, stable and robust strategies. The strategy *equalizerF* is the fourteens of the evolutionary competition, and does not confirm its success during the Exp3 experimentation. It does not seem as robust as the 3 winners.

4.16 As with all the experiments of this paper, containing a probabilistic strategy, this Exp4 experience is based on a tournament repeated 50 times between the involved strategies. For example, to check the stability of this result, here is the ranking obtained by the first five strategies after the first ten executions.

	Run1	Run2	Run3	Run4	Run5	Run6	Run7	Run8	Run9	Run10
mem2	1	1	1	1	2	1	1	1	1	1
gradual	2	2	2	2	1	2	4	2	2	2
spiteful	4	3	3	4	3	4	2	4	4	3
mem11_cCDDD-spite	3	4	4	3	4	3	3	3	3	4
soft_majo	5	5	5	5	5	5	5	5	5	5
mem11_cCDCD-tft	6	6	7	7	7	6	7	7	6	6
tit_for_tat	7	7	6	6	6	7	6	6	7	7
hard_tft	8	8	8	8	8	8	8	8	8	8
mem11_cCDCC	9	9	11	11	10	9	10	11	9	9
mem11_cCDDC-pavlov	10	10	10	10	11	11	9	9	10	10

4.17 This experiment shows that probabilistic strategies introduced by Press and Dyson are not good competitors (except for *equalizerF*, which is relatively efficient). This had already been noted in several papers (Hilbe et al. 2014, 2013; Stewart & Plotkin 2013; Adami & Hintze 2013, 2014; Dong et al. 2014; Szolnoki & Perc 2014). Press and Dyson strategies are designed to equal or beat each strategy encountered in a one-to-one game. The *all_d* strategy itself also is never beaten by another strategy, and is known to be catastrophic because she gets angry with everyone (except stupid non-reactive strategies) and therefore does not earn nearly point, especially in evolutionary competitions where only survive efficient strategies after a few generations. To win against any opponent is pretty easy, scoring points is more difficult! The right strategies in the prisoner's dilemma are not those who try to earn as many points than the opponent (such as equalizers) or require to earn more points than any other (as extortioners), these are the ones that encourage cooperation, know how to maintain it and even restore it if necessary after a sequence of unfortunate moves.

Complete classes alone

5.1 So as objectively confirm the results of the first experiments and also to identify other strategies that need to be added to our selection, we began to conduct competitions among all strategies coming from as large as possible complete classes. Our platform has allowed us to compete in tournament and evolutionary competitions families of 1,000 and even 6,000 strategies (our limit today). The results found are full of lessons.

memory(1,1)

5.2 The experiment `Exp5` starts with the results of the complete class of the 32 *memory(1,1)* strategies. It objectively shows that *spiteful*, *tit_for_tat* and *pavlov* are efficient strategies. We can see that the victory of *all_d* in the tournament cannot resist to the evolutionary competition.

	Tournament ranking			**Evolutionary ranking**	
1	mem11_dCCDD-alld	96000	1	mem11_cCDDD-spite	2126
	mem11_dCDDD-alld	96000	2	mem11_cCDCD-tft	701
	mem11_dDCDD-alld	96000	3	mem11_cCDDC-pavlov	214
	mem11_dDDDD-alld	96000	4	mem11_cCDCC	158
5	mem11_cDDDD	95952	5	mem11_dDDDD-alld	0
6	mem11_cCDDD-spite	95928	6	mem11_dDCDD-alld	0
7	mem11_dDDDC	94988	7	mem11_dCCDD-alld	0
8	mem11_dCDDC	92480	8	mem11_dCDDD-alld	0
9	mem11_dDCDC	87500	9	mem11_cDDDD	0
10	mem11_cDDDC	87450	10	mem11_dDDCD	0

5.3 When we consider complete classes we note the first plays (which do not depend on the past) in lowercases, and the other plays in uppercases.

memory(1,2)

5.4 The experiment `Exp6` concerns the *memory(1,2)* class (a move of my past, and two moves of the opponent's past) which contains 1024 strategies. To define a strategy for this class, we must choose what she plays in the first two moves (placed at the head of the genotype) and what she plays when the past was: `[c ; (c c)] [c ; (c d)] [c ; (d c)] [c ; (d d)] [d ; (c c)] [d ; (c d)] [d ; (d c)] [d ; (d d)]`.

	Tournament ranking			**Evolutionary ranking**	
1	mem12_ddCCDDDDDC	3397866	1	mem12_ccCDCDDCDD	20877
	mem12_ddCDDDDDDC	3397866	2	mem12_ccCDCDDDCD	8530
3	mem12_ddDCDDDDDC	3396868		mem12_ccCDCDDCCD	8530
	mem12_ddDDDDDDDC	3396868		mem12_ccCDCDCCCD	8530
5	mem12_ddDDCDDDDC	3333078		mem12_ccCDCDCDCD	8530
6	mem12_ddCDCDDDDC	3290142	6	mem12_ccCCCDDDDD	7451
7	mem12_ddDCCDDDDC	3273226	7	mem12_ccCDCDDDDD	6911
8	mem12_ddCCDCDDDC	3271234	8	mem12_ccCCCDDCDD	6750
	mem12_ddCDDCDDDC	3271234	9	mem12_ccCDCDCCDD	5248
10	mem12_ddDCDCDDDC	3270236	10	mem12_ccCDDDDCDD	1964

5.5 The winner is a strategy that plays *pavlov* except at the beginning where she plays `c,c` and, when she was betrayed twice, she betrays (unlike pavlov). We will name it *winner12*.

5.6 This *winner12* makes us think to a mixture as simple as possible of *tit_for_tat* and *spiteful*: She plays *tit_for_tat* unless she has been betrayed two times consecutively, in which case she always betrays (plays *all_d*). We will

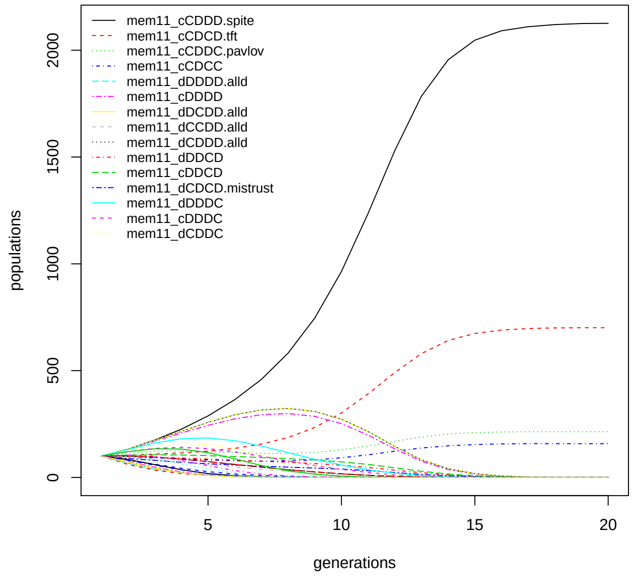

Figure 5: Evolutionary competition Exp5 involving the 32 *memory(1,1)* strategies.

call this new strategy *tft_spiteful* which to our knowledge has never been previously identified in any paper, despite its simplicity.

memory(2,1)

5.7 Exp7 experiment concerns the *memory(2,1)* complete class which contains 1024 strategies. The genotype is defined with the same principle as *memory(1,2)*.

	Tournament ranking				Evolutionary ranking	
1	mem21_dcDDDDDCDD	3180976		1	mem21_dcCDCDCDDD	50787
2	mem21_ddCCDDDDDC	3153294		2	mem21_dcCDCDCCDD	21680
	mem21_ddCDDDDDDC	3153294		3	mem21_dcCDCDCDCD	14716
4	mem21_ddDCDDDDDC	3152296		4	mem21_dcCDCDCCCD	3060
	mem21_ddDDDDDDDC	3152296		5	mem21_dcCDDDCCDD	2923
6	mem21_cdCCDDDDDC	3151798		6	mem21_dcCDCDCDDC	2169
	mem21_cdCDDDDDDC	3151798		7	mem21_dcCDCDCDCC	1629
8	mem21_cdDCDDDDDC	3150800		8	mem21_dcCDDDCDDD	1149
	mem21_cdDDDDDDDC	3150800		9	mem21_dcCDDDCDCD	962
10	mem21_dcCDDDDCDD	3077696		10	mem21_dcCDDDCCCD	577

5.8 The winner is a strategy that plays `tit_for_tat` except that it starts with `d,c`, and, when she betrayed twice and the other has nevertheless cooperated she reacts by a `d` (this is the only round that differentiates it from `tit_for_tat`). She exploits the kindness of the opponent. We will name it *winner21*.

5.9 The following slightly simpler and less provocative strategy (which is usually a quality) seemed interesting to us: she plays `cc` at the beginning and then plays *spiteful*. We call it *spiteful_cc* It is a kind of softened *spiteful*.

memory(1,2) + memory(2,1)

5.10 The Exp8 experiment shows a confrontation including the two complete classes: *memory(1,2)* and *memory(2,1)*. This leads to a set of 2,048 strategies.

5.11 Computing these results requires a 2,048 * 2,048 matrix to fill, so roughly 4 million meetings, and for each of them, 1,000 rounds. It also need for the evolutionary competition a population of 2,048 * 100 agents operating a thousand times.

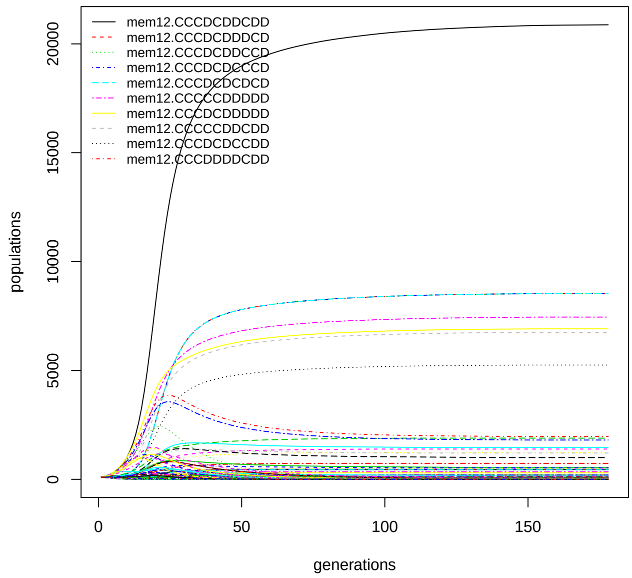

Figure 6: Evolutionary competition `Exp6` involving the 1,024 *memory(1,2)* strategies.

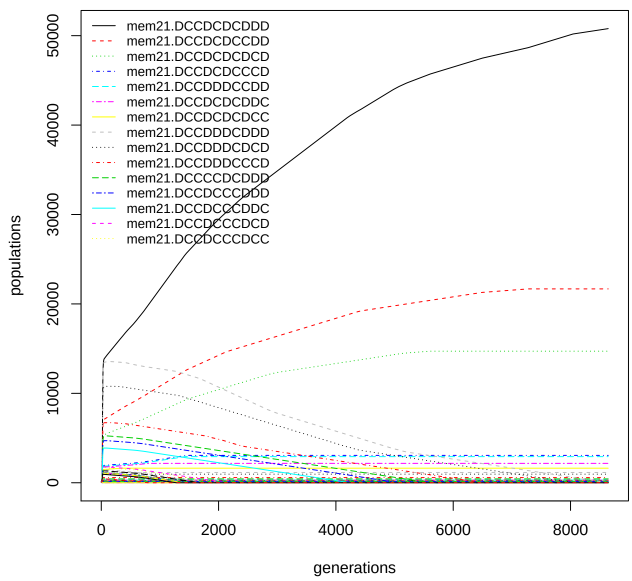

Figure 7: Evolutionary competition `Exp7` involving the 1,024 *memory(2,1)* strategies.

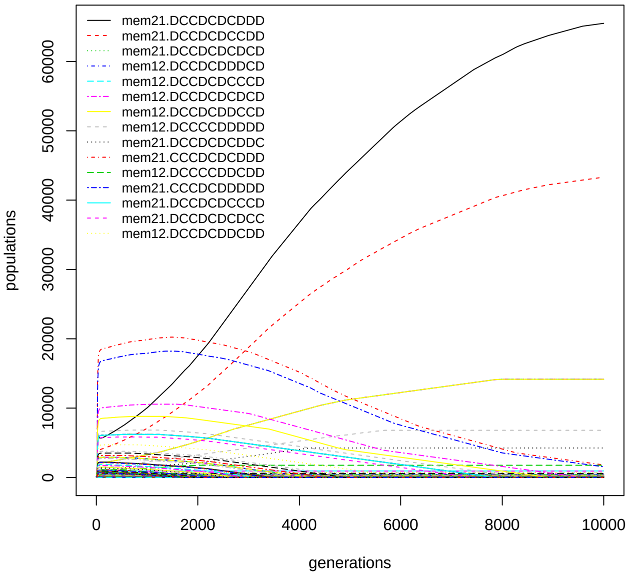

Figure 8: Evolutionary competition Exp8 involving 2,048 strategies built with the 1,024 *memory(1,2)* strategies + the 1,024 *memory(2,1)* strategies.

	Tournament ranking			Evolutionary ranking	
1	mem12_ddCCDDDDDC	6573258	1	mem21_dcCDCDCDDD	65503
	mem12_ddCDDDDDDC	6573258	2	mem21_dcCDCDCCDD	43308
3	mem12_ddDCDDDDDC	6572260	3	mem21_dcCDCDCDCD	14164
	mem12_ddDDDDDDDC	6572260		mem12_dcCDCDDDCD	14164
5	mem21_ddCCDDDDDC	6447758		mem12_dcCDCDCCCD	14164
	mem21_ddCDDDDDDC	6447758		mem12_dcCDCDCDCD	14164
7	mem21_ddDCDDDDDC	6446760		mem12_dcCDCDDCCD	14164
	mem21_ddDDDDDDDC	6446760	8	mem12_dcCCCDDDDD	6802
9	mem12_ddDDCDDDDC	6422478	9	mem21_dcCDCDCDDC	4247
10	mem12_ddCCDCDDDC	6360918	10	mem21_ccCDCDCDDD	1803

5.12 It is remarkable that the winner is *winner21*. It remains to be seen whether the 4 new strategies we have just introduced are really robust, and how they are ranked when confronted to the best previously identified strategies.

Same experiments with the 4 new strategies

6.1 We take once again the first 4 experiments done in Sections 3 and 4, each time adding our four new strategies, which allows us to evaluate both the robustness of former winners and put them in competition with the new four.

17 basic + 4 new strategies

6.2 The experiment Exp9 involves the 17 basic strategies like in Exp1 (Section 3.7) with the four new strategies dis-

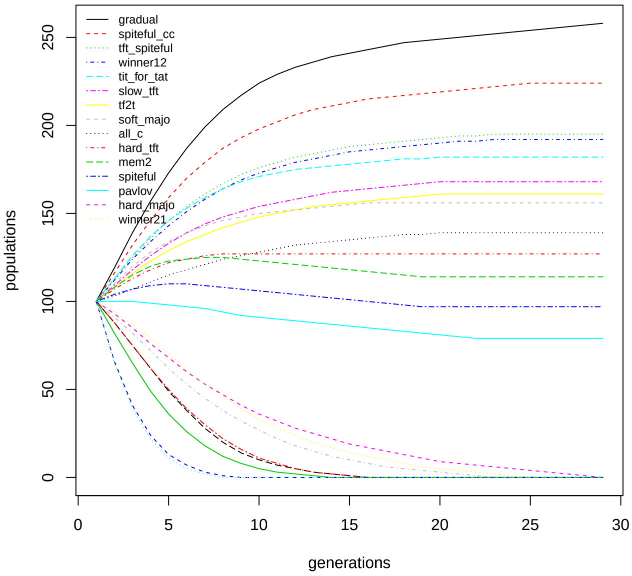

Figure 9: Evolutionary competition Exp9 involving the 17 basic strategies + the 4 new discovered strategies.

covered thanks to the complete classes experiments (Sections 5.4 and 5.7).

Tournament ranking			Evolutionary ranking		
1	gradual	60823	1	gradual	258
2	spiteful_cc	58981	2	spiteful_cc	224
3	tit_for_tat	57661	3	tft_spiteful	195
4	tft_spiteful	57486	4	winner12	192
5	winner12	57072	5	tit_for_tat	182
6	soft_majo	56330	6	slow_tft	168
7	slow_tft	55821	7	tf2t	161
8	tf2t	55156	8	soft_majo	156
9	mem2	55014	9	all_c	139
10	hard_tft	54679	10	hard_tft	127

6.3 It is remarkable that three among the four new introduced strategies are in the four first evolutionary ranking.

6.4 It appears here that *mem2* is not a robust strategy. She is the 9^{th} in tournament and is not even in the top 10 in the ecological competition.

30 + 4 new strategies

6.5 Exp10 studies the 30 deterministic and probabilistic basic strategies like in Exp2 (Section 3.10) with the four new strategies discovered thanks to the complete classes experiments (Section 5.4 and 5.7). This leads to a set of 34 strategies.

Tournament ranking			Evolutionary ranking		
1	soft_majo	4233818	1	gradual	391
2	gradual	4161655	2	spiteful_cc	330
3	equalizerF	4106990	3	soft_majo	291
4	all_c	4050351	4	tft_spiteful	276
5	spiteful_cc	4049049	5	equalizerF	274
6	tft_spiteful	3975187	6	winner12	268
7	tit_for_tat	3949101	7	all_c	252
8	winner12	3929559	8	tit_for_tat	250
9	slow_tft	3862035	9	slow_tft	222
10	mem2	3852253	10	tf2t	211

6.6 The strategy *gradual* wins, and strangely, *all_c* is the seventh, but the three new introduced strategies (*spiteful_cc*, *winner12*, *tft_spiteful*) are among the 10 best.

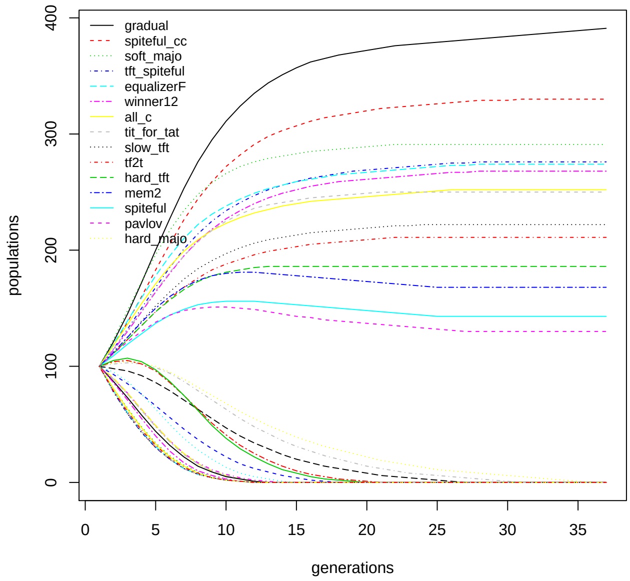

Figure 10: Evolutionary competition Exp10 involving the 30 deterministic and probabilistic initial strategies + the 4 new discovered strategies.

6.7 For example, to check the stability of the Exp10 result, here is the ranking obtained by the first five strategies after the first ten executions.

	Run1	Run2	Run3	Run4	Run5	Run6	Run7	Run8	Run9	Run10
gradual	1	1	1	1	1	1	1	1	1	1
spiteful_cc	2	2	2	2	3	2	2	2	2	2
soft_majo	3	5	3	3	4	3	3	3	1	3
tft_spiteful	5	4	4	4	5	4	4	4	5	4
equalizerF	4	3	5	> 5	2	5	5	5	4	5

6.8 Note that in this table the extreme stability of the beginning of ranking. Except from the run4, the first five strategies are always the same. Of course, lower one goes in these rankings, more there are permutations, but the first five remain the same.

All deterministic + 4 new strategies

6.9 For the Exp11 we take all the deterministic strategies obtained with the 17 initial basic strategies and the *memory(1,1)* complete classes, thus $17 + 32$ like in Exp3 (Section 4.10) with the four new strategies discovered thanks to the complete classes experiments (Sections 5.4 and 5.7). This leads to a set of 53 strategies.

	Tournament ranking			**Evolutionary ranking**	
1	spiteful_cc	152873	1	spiteful_cc	528
2	gradual	150685	2	gradual	520
3	winner12	149466	3	winner12	467
4	spiteful	148934	4	tft_spiteful	438
	mem11_cCDDD-spite	148934	5	mem2	345
6	mem2	146936	6	mem11_cCDDD-spite	343
7	tft_spiteful	144068		spiteful	343
8	tit_for_tat	132809	8	tit_for_tat	286
	mem11_cCDCD-tft	132809		mem11_cCDCD-tft	286
10	pavlov	132712	10	soft_majo	241

6.10 This time, the four winners are exactly the same as in Exp9 but not exactly in the same order. This result shows the robustness of these four strategies.

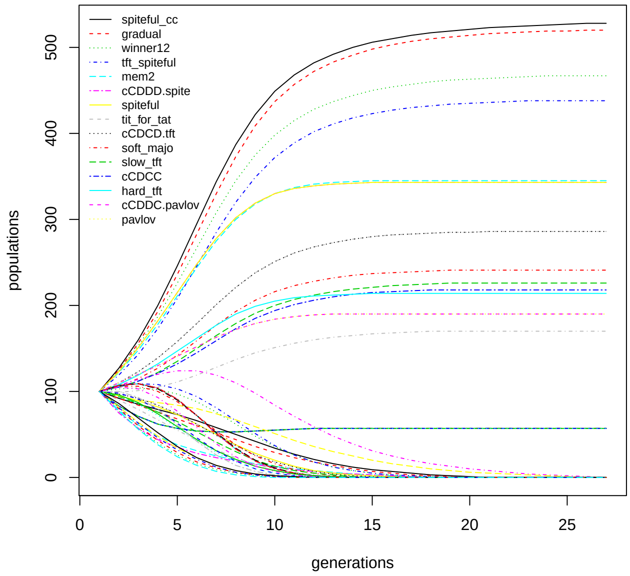

Figure 11: Evolutionary competition Exp11 involving 53 strategies built with the 17 basic strategies + the 32 *memory(1,1)* strategies + the 4 new discovered strategies.

All deterministic and probabilistic

6.11 Exp12 is built with all the basic deterministic strategies obtained with the 17 initial basic strategies and the *memory(1,1)* complete class added with the 13 probabilistic strategies like in Exp4 (Section 4.14) with the four new strategies discovered thanks to the complete classes experiments (Sections 5.4 and 5.7). This leads to a set of 66 strategies.

	Tournament ranking			**Evolutionary ranking**	
1	spiteful_cc	8746187	1	spiteful_cc	641
2	gradual	8657237	2	gradual	622
3	mem11_cCDDD-spite	8550396	3	winner12	543
4	winner12	8549848	4	tft_spiteful	518
5	spiteful	8545930	5	mem11_cCDDD-spite	421
6	mem2	8447939		spiteful	421
7	tft_spiteful	8305044		mem2	421
8	pavlov	7898010	8	soft_majo	341
9	mem11_cCDDC-pavlov	7896930	9	tit_for_tat	316
10	soft_majo	7820688	10	mem11_cCDCD-tft	315

6.12 Once again, the same four strategies win this competition. This confirms the results obtained during Exp1 to Exp8 experiments. *winner21* is only 16th in this ranking.

Stability and Robustness of the Results

7.1 To test the stability of these results, we have built a set of five experiments. The first one test if probabilistic strategies makes the ranking unstable. The second test measures the effects of the length of the meetings. The third test verifies that the changes of coefficients in the payoff matrix have any effect. The last test ensures that even when taking strategies that have a longer memory and using diversified strategies, the results are always stable.

Test with respect to probabilistic strategies

7.2 In previous experience Exp12, scores are obtained by averaging over 50 rounds to ensure stability. To see in detail the influence of probabilistic strategies we point out, 10 classifications obtained without making any average.

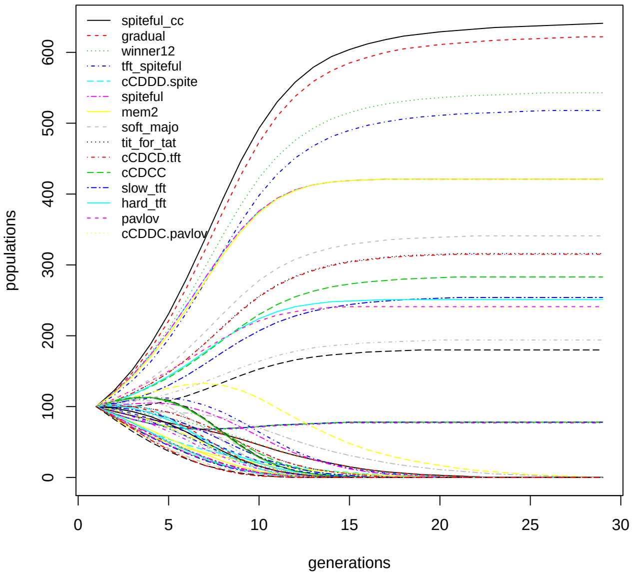

Figure 12: Evolutionary competition Exp12 involving 66 strategies built with the 17 initial basic strategies + the 32 *memory(1,1)* strategies + the 13 probabilistic strategies + the 4 new discovered strategies.

	Run1	Run2	Run3	Run4	Run5	Run6	Run7	Run8	Run9	Run10
spiteful_cc	1	1	1	1	1	1	1	1	1	1
gradual	2	3	2	2	2	2	2	3	2	2
winner12	3	2	3	3	3	3	3	4	4	4
tft_spiteful	4	4	4	4	4	4	4	2	3	3
mem11_cCDDD-spite	7	5	6	5	5	6	5	6	6	6
mem2	6	6	5	6	7	5	6	5	5	7
spiteful	5	7	7	7	6	7	7	7	7	5
soft_majo	10	10	8	8	9	9	8	8	8	8
mem11_cCDCD-tft	8	9	9	10	8	8	9	10	10	9
tit_for_tat	9	8	10	9	10	10	10	9	9	10

7.3 We can see that the first ten strategies are always the same. Only their ranking changes.

Test with respect to meetings lengths

7.4 Previous experiences were made with 1,000 rounds by meeting. We are now testing whether the length of the meetings influences many rankings.

name	length 10	length 20	length 50	length 100
spiteful_cc	1	1	1	1
gradual	10	7	3	2
winner12	6	2	4	3
tft_spiteful	5	3	2	4
mem11_cCDDD-spite	4	5	6	5
mem2	2	4	5	6
spiteful	3	6	7	7
soft_majo	9	10	9	8
tit_for_tat	8	9	8	9
mem11_cCDCD-tft	7	8	10	10

7.5 One can see that, when the length of the meeting is greater that 10 rounds, then the first 10 strategies stay the same. Just their ranking changes. From a length of 60, nothing changes in the ranking of the first 10. We note that shorter the meetings are, more `mem2` is favoured and less `gradual` is disadvantaged. This results shows clearly that the qualities of `gradual` require a certain length of meeting.

Test with respect to the payoff matrix

7.6 In this section we change the coefficients of the experience Exp12 (Section 6.11) by transforming (5, 3, 1, 0) to (2, 1, 0, 0) in the matrix of gains, to test the stability relative to earnings, while remaining under the classic dilemma of inequality. These coefficients corresponds to the British TV show on ITV Networks called *"Golden Balls, Split or Steal"*. This experiment have been repeated fifty times with 1000 rounds meetings.

	Tournament ranking			**Evolutionary ranking**	
1	spiteful_cc	2723515	1	spiteful_cc	649
2	winner12	2702022	2	gradual	589
3	gradual	2695240	3	winner12	578
4	mem11_cCDDD-spite	2625792	4	tft_spiteful	568
5	spiteful	2625237	5	mem11_cCDDD-spite	423
6	mem11_cCDDC-pavlov	2614751	6	spiteful	422
7	pavlov	2614260	7	mem2	404
8	tft_spiteful	2608724	8	mem11_cCDCC	364
9	mem11_dDCDC	2602493	9	tit_for_tat	341
10	mem11_dCDDC	2597292		mem11_cCDCD-tft	341

7.7 These results have to be compared with those of Exp12 (see Section 6.11) which are quite the same.

Test of independence

7.8 To test if the four new strategies are individually efficient, that is their good results do not depend from the others, we make compete each of the 17 + 4 strategies one of one, with the set of 1024 *memory(1,2)*. In each of these 21 experiments involving 1025 strategies, we measure this time the rank of the added strategy.

strategy	rank
tft_spiteful	1
winner12	1
spiteful_cc	2
gradual	10
tit_for_tat	13
slow_tft	14
mem2	19
tf2t	28
winner21	32
spiteful	34
soft_majo	37
hard_tft	67
mistrust	95
pavlov	95
hard_majo	167
per_cd	172
per_ddc	294
prober	351
all_d	390
per_ccd	564
all_c	919

7.9 One can see on these results that if we just add *tft_spiteful* to the set of 1024 *memory(1,2)* strategies, it finishes first. This is also the case obviously for *winner12*. In the same way, *spiteful_cc* finishes second. On the other hand *tit_for_tat* finishes only 13^{th}. Again, we find that among the 4 added strategies, 3 of them are really excellent.

Test with a rich soup

7.10 As it is impossible to run large complete classes (*memory(2,2)* contains for example 262,144 strategies), one example have been obtained by taking randomly 1,250 strategies from `memory(2,2)` + 1,250 strategies from *memory(3,3)* + 1,250 strategies from *memory(4,4)* + 1,250 strategies from *memory(5,5)* with the now famous 17+4.

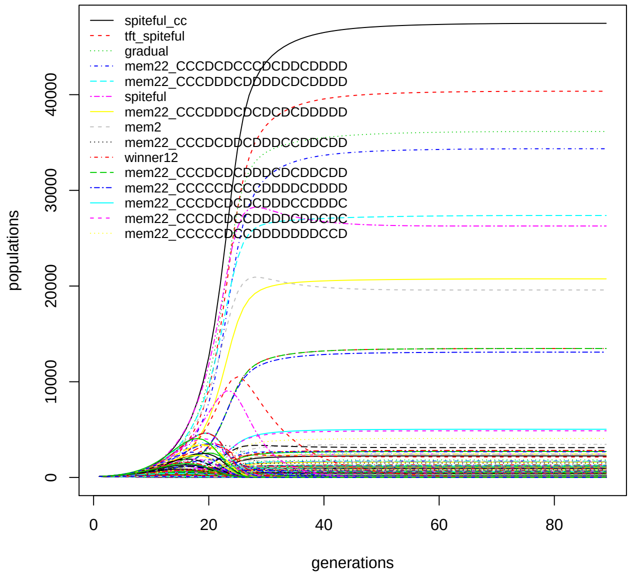

Figure 13: Evolutionary competition involving 5021 strategies built with the 17 initial basic strategies + the 4 new discovered + 1250 of each *memory(X,X)* with X in 2,3,4,5 chosen randomly. This figure comes from one of the twenty cumulative experiences used in Section 7.10

This set contains then 5,021 strategies. This experiment is run twenty times to be able to compute relevant rank average and standard deviation.

strategy	rank avg	sd
tft_spiteful	1.6	0.9
spiteful_cc	2.4	0.76
winner12	8.45	2.64
gradual	9.25	4.875
mem2	18.2	19.94
spiteful	22.2	24.94
tit_for_tat	28.65	13.215
slow_tft	30.7	6.4
hard_tft	30.9	21.13
soft_majo	37.25	12.9
tf2t	84.3	12.27
winner21	128.75	32.125
mistrust	157.75	18.875
hard_majo	167.8	13.5
pavlov	237.2	93.44
prober	309.85	39.62
all_d	315.85	34.35
per_ddc	621.95	104.05
per_cd	1122.5	318.65
per_ccd	4430.15	70.865
all_c	4995.65	7.555

7.11 This test illustrates once again that three of the four (*spiteful_cc*, *tft_spiteful* and *winner12*) new introduced strategies are in the top (1,2 and 3). One can note also the great robustness of *gradual* who finished fourth of this huge experiment.

Test of evolutionary stability

7.12 In order to add a robustness test to the strategies identified, we conducted a series of experiments to test their stability against invasions of different types.

7.13 We are interested in the following 10 strategies which are the best knows strategies resulting from our experiments: *tft_spiteful* , *spiteful_cc*, *winner12* , *gradual*, *mem2* , *spiteful*, *tit_for_tat*, *slow_tft*, *hard_tft*, *soft_majo*.

7.14 In turn, we take 10,000 copies of *all_d* and 10,000 copies of one of the 10 previously mentioned strategies that come together in an evolutionary competition. In each case, *all_d* is quickly and totally eliminated.

7.15 We then changed the proportion (10,000 vs 10,000) by gradually decreasing the numbers of each of the strategies studied. *all_d* is always eliminated, except when the number of the strategy added is less than 75 copies. For example, 10,000 *all_d* are eliminated by 100 *winner12*, but are not eliminated by 60.

7.16 The same experiment has been performed by replacing *all_d* by the *random* strategy. This time the *soft_majo* strategy proves to be weaker: the switching is done at approximately 500 while for the others the switching is at approximately 200 which confirms the robustness to the invasion of our 10 selected strategies.

7.17 Not only do these 10 strategies not let themselves be invaded by others, they invade the others, even when their starting population are much lower.

7.18 More in-depth methods for studying evolutionary stability can be envisaged using methods described in Ficici & Pollack (2003); Ficici et al. (2005).

Conclusion

8.1 According to the state of the art, in the first part of this paper we have collected the most well-known interesting strategies. Then we have used the systematic and objective complete classes method to evaluate them. These experiments led us to identify new efficient and robust strategies, and more than that, a general scheme to find new ones. The four new strategies are actually successful strategies, even if *winner21* seems less robust. Although detected by calculating in special environments the three new robust strategies (*spiteful_cc*, *winner12*, *tft_spiteful*) remain excellent even in other environments unrelated to that of their "birth". The method of complete classes is clearly an efficient method to identify robust winners.

8.2 At this time, we consider, according with the final ranking in Section 7.10 that the best actual strategies in the IPD are in order

tft_spiteful* , *spiteful_cc*, *winner12*, *gradual
mem2*, *spiteful*, *tit_for_tat*, *slow_tft*, *hard_tft*, *soft_majo

8.3 The two best strategies come from this paper. We encourage the community to take systematically into account these new strategies in their future studies.

8.4 We note that these are almost all mixtures of two basic strategies: *tit_for_tat* and *spiteful*. This suggests that *tit_for_tat* is not severe enough, that *spiteful* is a little too much severe and that finding ways to build hybrids of these two strategies is certainly what gives the best and most robust results.

8.5 We also note that using information about the past beyond the last move is helpful. Among the eight strategies that our tests put in the head of ranking some of them use the past from the beginning (*gradual* and *soft_majo*) and all the others use (except *equalizer-F*) two moves of the past or a little more. The memory also seems useful to play well (confirming the results of (LI & Kendall 2013; Moreira et al. 2013)).

8.6 A promising way to find other efficient strategies is probably to carefully study larger complete classes, to identify the best and check their robustness. The lessons learned from these experiments generally concern many multiagent systems where strategies and behaviours are needed.

References

Adami, C. & Hintze, A. (2013). Evolutionary instability of zero-determinant strategies demonstrates that winning is not everything. *Nature Communications*, *4*(2193)

Adami, C. & Hintze, A. (2014). Corrigendum: Evolutionary instability of zero-determinant strategies demonstrates that winning is not everything. *Nature Communications*, *4*, 2193

Axelrod, R. M. (2006). *The Evolution of Cooperation*. New York, NY: Basic Books

Beaufils, B., Delahaye, J.-P. & Mathieu, P. (1996). Our meeting with gradual, a good strategy for the iterated prisoner's dilemma. In *Proceedings of the Fifth International Workshop on the Synthesis and Simulation of Living Systems (ALIFE'5)*, (pp. 202–209). Boston, MA: The MIT Press/Bradford Books

Beaufils, B., Delahaye, J.-P. & Mathieu, P. (1998). Complete classes of strategies for the classical iterated prisoner's dilemma. In *Evolutionary Programming VII (EP'7)*, vol. 1447 of *Lecture Notes in Computer Science*, (pp. 33–41). Berlin/Heidelberg: Springer

Beaufils, B. & Mathieu, P. (2006). Cheating is not playing: Methodological issues of computational game theory. In *Proceedings of the 17th European Conference on Artificial Intelligence (ECAI'06)*, (pp. 185–189). Amsterdam: IOS Press

Delahaye, J.-P., Mathieu, P. & Beaufils, B. (2000). The iterated lift dilemma. In *Computational Conflicts : Conflict Modeling for Distributed Intelligent Systems*, (pp. 203–223). Berlin/Heidelberg: Springer

Dong, H., Zhi-Hai, R. & Tao, Z. (2014). Zero-determinant strategy: An underway revolution in game theory. *Chinese Physics B*, *23*(7), 078905

Ficici, S. & Pollack, J. (2003). A game-theoretic memory mechanism for coevolution. In *Genetic and Evolutionary Computation – GECCO 2003*, (pp. 203–203). Berlin/Heidelberg: Springer

Ficici, S. G., Melnik, O. & Pollack, J. B. (2005). A game-theoretic and dynamical-systems analysis of selection methods in coevolution. *IEEE Transactions on Evolutionary Computation*, *9*(6), 580–602

Hilbe, C., Nowak, M. A. & Sigmund, K. (2013). Evolution of extortion in Iterated Prisoner's Dilemma games. *Proceedings of the National Academy of Sciences*, *110*(17), 6913–6918

Hilbe, C., Röhl, T. & Milinski, M. (2014). Extortion subdues human players but is finally punished in the Prisoner's Dilemma. *Nature Communications*, *5*, 3976

Kendall, G., Yao, X. & Chong, S. Y. (2007). *The Iterated Prisoners' Dilemma: 20 Years on*. Singapore: World Scientific Publishing Co.

Li, J., Hingston, P. & Kendall, G. (2011). Engineering design of strategies for winning iterated prisoner's dilemma competitions. *Computational Intelligence and AI in Games, IEEE Transactions on*, *3*(4), 348–360

LI, J. & Kendall, G. (2013). The effect of memory size on the evolutionary stability of strategies in iterated prisoner's dilemma. *Evolutionary Computation, IEEE Transactions on*, *PP*(99), 1–8

Lorberbaum, J. (1994). No strategy is evolutionarily stable in the repeated prisoner's dilemma. *Journal of Theoretical Biology*, *168*(2), 117–130

Lorberbaum, J. P., Bohning, D. E., Shastri, A. & Sine, L. E. (2002). Are there really no evolutionarily stable strategies in the iterated prisoner's dilemma? *Journal of Theoretical Biology*, *214*(2), 155–169

Mathieu, P., Beaufils, B. & Delahaye, J.-P. (1999). Studies on dynamics in the classical iterated prisoner's dilemma with few strategies: Is there any chaos in the pure dilemma ? In *Proceedings of the 4th european conference on Artificial Evolution (AE'99)*, vol. 1829 of *Lecture Notes in Computer Science*, (pp. 177–190). Berlin/Heidelberg: Springer

Mathieu, P. & Delahaye, J.-P. (2015). New winning strategies for the iterated prisoner's dilemma. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, (pp. 1665–1666). International Foundation for Autonomous Agents and Multiagent Systems

Mittal, S. & Deb, K. (2009). Optimal strategies of the iterated prisoner's dilemma problem for multiple conflicting objectives. *Evolutionary Computation, IEEE Transactions on*, *13*(3), 554–565

Moreira, J., Vukov, J., Sousa, C., Santos, F. C., d'Almeida, A. F., Santos, M. D. & M., P. J. (2013). Individual memory and the emergence of cooperation. *Animal Behaviour*, *85*(1), 233 – 239

O'Riordan, C. (2000). A forgiving strategy for the iterated prisoner's dilemma. *Journal of Artificial Societies and Social Simulation*, *3*(4), 3

Poundstone, W. (1992). *Prisoner's Dilemma: John von Neuman, Game Theory, and the Puzzle of the Bomb*. New York, NY: Doubleday

Press, W. H. & Dyson, F. J. (2012). Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent. *Proceedings of the National Academy of Sciences*, *109*(26), 10409–10413

Rapoport, A. & Chammah, A. (1965). *Prisoner's Dilemma: A Study in Conflict and Cooperation*. Ann Arbor, MI: University of Michigan Press

Sigmund, K. (2010). *The Calculus of Selfishness*. Princeton, NJ: Princeton University Press

Stewart, A. J. & Plotkin, J. B. (2013). From extortion to generosity, evolution in the iterated prisoner's dilemma. *Proceedings of the National Academy of Sciences*, *110*(38), 15348–15353

Szolnoki, A. & Perc, M. (2014). Defection and extortion as unexpected catalysts of unconditional cooperation in structured populations. *Scientific Reports*, *4*, 5496

Tzafestas, E. (2000). Toward adaptive cooperative behavior. In *Proceedings of the Simulation of Adaptive Behavior Conference*. Paris

Wedekind, C. & Milinski, M. (1996). Human cooperation in the simultaneous and the alternating prisoner's dilemma: Pavlov versus generous tit-for-tat. *Proceedings of the National Academy of Sciences*, *93*(7), 2686–2689

Wellman, M. P. (2006). Methods for empirical game-theoretic analysis. In *Proceedings of the National Conference on Artificial Intelligence*, vol. 21, (p. 1552). Cambridge, MA: MIT Press