The Iterated Lift Dilemma
or
How to Establish Meta-cooperation with your Opponent*

Jean-Paul Delahaye and Philippe Mathieu
Laboratoire d'Informatique Fondamentale de Lille
U.A. 369 du C.N.R.S., University of Lille I,
59655 Villeneuve d'Ascq Cedex. FRANCE.
e-mail: {delahaye,mathieu}@lifl.fr

Abstract

A very small change in Iterated Prisoner's Dilemma (I.P.D.) payoff matrix leads to an iterated game called Iterated Lift Dilemma[1] which properties are very different than classical I.P.D. The following points are to be noted: (i) two levels of cooperation are now possible, the best one needs a difficult coordination between considered strategies; (ii) only probabilistic strategies can make a hight score when the play against themselves; (iii) more complex dynamics can appear (the "edge of chaos" ?) as soon as three strategies are confronted. Our idea, already defended about the classical IPD, is that, in spite of the model simplicity you can obtain many complex phenomena: it is not true that to be good, a strategy must be simple. Building good strategies for the Iterated Lift Dilemma is then much more difficult than for the classical game.

1 Introduction

An *iterated game* is a game with two players A and B (also called *strategies*) which play an unknown finite number of rounds. On each round, each player chooses between the two actions c (for *cooperate*) and d (for *defect*). A round where the player A plays c and the player B plays d is noted [c,d]; a round where the player A plays c and the player B plays c is noted [c,c]; a round where the player A plays d and the player B plays d is noted [d,d].
When the players play the round number n they play simultaneously with taking account the game history (that is all the preceding choices they both have made at all rounds i with $i < n$).
A player can also play a round randomly; in such a case we say that it is a probabilistic strategy.
The average length of each game must be large enough (> 10) to permit interesting phenomena and to obtain robust results.
The results of the player's choices are quantified by a payoff matrix. The parameter R is the reward for mutual cooperation (round [c,c]), T is the temptation to defect against a cooperation opponent who then gets the sucker's payoff S (round [d,c]). In case of mutual defection both gets the punishment P (round [d,d]).
Several interesting confrontations can be studied:

- There are *single confrontations* (one strategy against another one). At the end of the confrontation (for example after 100 rounds) points obtained by each player are added, the winner is the player which has the greatest score.

*This work has been partially supported by the Contrat-Plan Etat-Region GANYMEDE II
[1]The term "Lift" comes from the French expression "renvoi d'ascenceur" which means "I help you this time, you will help me next time"

- There are *round-robin tournaments*. We take k strategies, each one playing against all the others (including itself). Points of confrontations are added, the winner is the player which has the greatest score.

- There are *ecological evolutions*. For example we consider 3 strategies A,B,C; we start with a population of one hundred of A strategies, one hundred of B strategies and one hundred of C strategies, that is 300 entities (this defines the first generation); a round-robin tournament is computed (each entity plays against the 299 other entities); scores are computed for each strategy (sum of all score entities of the same kind); a new population is then computed for each strategy which is proportional the score it obtained. For simplicity we consider that the total population is constant (here 300). This defines the second generation. This computation is repeated again until populations become stable.

We could leave the total population to increase but this is not really significant since we are interested in relative strategies range. Thus our choice is to maintain the global population stable. In *ecological evolutions*, good strategies proliferate and replace bad ones which completely disappear. To be a good strategy in an *ecological evolution* you need not only to be good in the round-robin competition but also during all the time and especially when bad ones disappear. Persistent strategies in *ecological computation* are really robust ones.
In the classical Iterated Prisoner's Dilemma we often choose the following parameters:

$$S = 0 \qquad P = 1 \qquad R = 3 \qquad T = 5$$

which verify:

$$S < P < R < T \text{ and } (S + T)/2 < R$$

This game has been found to be a very good way of studying cooperation and evolution of cooperation and thus a sort of theory of cooperation based upon reciprocity has been set in a wide literature, such as in [1, 3, 4]. The experimental studies of the IPD and its strategies need a lot of time computation and thus with the progress of computers, a lot of computer-scientists and mathematicians have studied it as they have been able to use specific methods, like genetic algorithms, on it, see [2, 5, 9, 15, 16, 21, 22, 24, 25].
As cooperation is a topic of continuing interest for the social, zoological and biological sciences, a lot of works in those different fields have been made on the IPD: [6, 8, 12, 13, 14, 17, 18, 20, 23, 19].

In this paper we study the consequences of the second equality inversion:

$$(S + T)/2 > R$$

For example we will study the following parameters:

$$S = 0 \qquad P = 1 \qquad R = 3 \qquad T = 8$$

This small change on the classical game entails many surprising consequences. It becomes now more interesting to be according with its rival for playing [c,d] [d,c] [c,d] [d,c] ... (which gives an average of 4 points each round for each player) than playing [c,c] [c,c] [c,c] ... (which only gives an average of 3 points each round for each player).
It is always a dilemma: due to the first classical inequality that has not been changed, the collective interest contradicts the selfish interest. To maximize reward needs a subtle agreement
As for the classical IPD it is easy:

- to find cycles (A wins against B, B wins against C, C wins against A);

- to find infinite hierarchical classification (A1 wins against A0, A2 wins against A1, etc.);

- to show that there is no strategy which plays optimally (that is obtains the best possible score) against every other opponent;

- to show that DEFECT (which always defects) never looses against any other strategy but earns very few points each time (it wins against its opponents but each game it plays costs a lot for it);
- to show that TIT-FOR-TAT (which cooperates on the first move and then plays what its opponent played on the previous move) never looses more than 8 points against any other strategy.

With our new parameters the classical game analysis must be revisited. There are now two cooperation levels:

- the basic level (which looks like a "non aggression pact"): to play always [c,c], which gives an average reward of 3 points by round for each player;
- the upper level (or *meta-cooperation* level): to agree with the opponent to win and loose alternatively, that is to play [c,d] [d,c] [c,d] [d,c] ..., which gives an average reward of 4 points by round for each player.

To have success in meta-cooperation you must play in *opposite phase* [c,d] [d,c] [c,d] [d,c] ... which is difficult because it needs some *coordination* and a great risk of loss for the player which plays c first (it could wait reciprocity for a long time !)
Other hight-level cooperations are also possible: for example, playing [c,d] [c,d] [d,c] [d,c] [c,d] [c,d] [d,c] [d,c] ... (periodicity 4). Such meta-cooperations needs much more *coordination* and *confidence.*
More complex synchronization schemes are now possible but they are much more difficult to establish and to maintain.
Note that you cannot have a preliminary agreement with your opponent since choices are simultaneous. Other models are possible in which players make their choices alternatively. We do not study these models here [12].
The "Lift Dilemma" does not take in account all the synchronization and meta-cooperation problems, but it is a simple and clean model, and thus allows us to increase our general understanding of cooperation and reciprocity. As we will see, this model is in fact astonishingly more subtle than the classical IPD.

2 Real examples of Iterated Lift Dilemma

Here are examples of situations which are better described with the Iterated Lift Dilemma than with the IPD.

2.1 Elections with two candidates of the same party

Two members of the same political party X want to be candidate to the next local election. Of course there are also other candidates from other parties. Here are the different possible situations:

- `[c,c]`. The two candidates of party X take place in the elections but stay fair-play (that is, they will not mutually discredit themselves or trying to injure their respective reputations).

  Chances to be elected are equally shared among them. The party does not loose any votes. We evaluate that each candidate has 30% chances to be elected.
- `[d,d]`. The two candidates takes places in an aggressive competition which makes damage to their reputations. Their fight afraid some electors and the party globally looses votes. We evaluate now that each candidate has 10% chances to be elected.

- `[d,c]`. Only one of them is aggressive while the other stays quiet (or even does not really want to win or calls electors to vote for his colleague). Electors are not afraid (amused ?)

  Due to the fact that all votes for the party X are now concentrated on the same candidate We evaluate that he has now 80% chances to be elected while the kind one does not have any chances to be elected (0%).

In such a case, if there are many elections, it is obvious that the two candidates have to alternatively give way to their colleague (*meta-cooperation*). This is clearly the best global behavior.
Of course in real life, the candidate who gives way to the other hopping to have a feedback takes a great risk (political change, new candidates, defection of the other).
The number of rounds in such a game seems to be rather limited, but the feedback between A and B can already be realized in a different form in future elections, then the total number of rounds can easily increases until 5 or 10.

2.2 Collaborator's recruitment cession

During collaborator's recruitment cession (in research centers, university, or private company) explicit or implicit agreements between two different teams or sectors against the others are common. They often work in this way: "This time I will not defend too much my candidate and I help you to support your's. Next time you will support my future candidate".
A round [c,c] is a cession where each team kindly defends its candidate, a round [d,d] is a cession where the two teams fight roughly for the recruitment of its candidate (with a great risk to see the candidate of a third team to be chosen), a round [d,c] is a cession where one of the two teams leaves its chances for the other hopping a feedback the next time.

2.3 Sale by auction for art objects

If two collectors are in the same room where the objects they want are presented, it is better for them to agree themselves to buy alternatively the objects, instead of out-bidding mutually which leads a great global increase of each price. Their agreement is the following one: "I stay quiet during the auction of this object but please stay quiet for the next one, by this way we will earn our money".
A round [c,c] is an auction where we try quietly to buy the art object. A round [d,d] is an auction where we wildly out-bid to have the object. A round [d,c] is an auction where one of the two buyers abstain from saying anything hopping a feedback next time.

2.4 The two music amateur neighbors

The music amateurs neighbors have also to alternate their listening periods if they want to ear their music in good conditions.

[c,d]:	I ear my music and I am not disturbed by yours. I obtain a satisfaction of 8 "pleasure points".
[d,c]:	You ear your music and I don't ear mine. It counts for 0 "pleasure points" for me.
[d,d]:	I am trying to ear my music but I ear simultaneously yours. I just obtain 1 "pleasure point" but you also !
[c,c]:	We renounce together to ear music today. This silence counts for 3 "pleasure points".

2.5 The access to an indivisible thing which is periodically available

This case is a generalization of the previous cases.

c:	Trying to catch this thing in respect to fairness or non aggression pact.
d:	Trying to catch this thing without restraint, defecting any implicit or explicit agreement defining fight rules.

A round [d,c] corresponds then to a case where one of the players submits to the other player's authority. This avoid the fight to degenerate violently.

In the animal world, you can frequently see some fights not really violent taking place between two animals of the same species, until one of the opponents gives up and shows its defection with a conventional sign. This shows clearly that such a situation is a kind of *Lift Dilemma*.
A round [c,c] corresponds to a fight not really violent during a long time (for example between to males which desire the same female). A round [d,d] corresponds to a violent fight which can result in hard wounds or possibly in the death of one of the opponents.
In animal fight, we can rarely see a meta-cooperation level, but more frequently we see a hierarchical situation with the consequence that the first winner always win in the following confrontations. We will see in the studies on homogeneous populations that this situation corresponds to the choice of a *collectively rational* strategy (defined as a strategy which is able to obtain a maximum score against itself even if the rewards are not equally distributed between its representatives).

3 What would be a good strategy ?

Before giving mathematical results and reporting computer experiments, it is interesting to elaborate an *a priori* analysis of the game. The following points are to be noted:

- As for the classical IPD, this game is a non-zero sum game (the total score distributed among the players depends on the actions chosen). Solidarity between the players comes from the game rules. This means that to have success players must be able to establish cooperation (or meta-cooperation) to obtain the maximum global score.
- It is really a dilemma: the only Nash equilibrium (if one of the player changes its position in any way, it will loose points) is a round [d,d]. Of course, the only non-dominated strategy is DEFECT (which always defect). By construction, the game is more difficult than the classical IPD due to the existence of two levels of cooperation and also to the hight quality of coordination needed to obtain a maximum global score (that is meta-cooperation).
- There is a kind of paradox in that to reach the meta-cooperation level one of the two players must defect first: by this way there is a risk to afraid an anxious opponent which will think that you don't want to cooperate. For example the SPITE strategy (which cooperates until you cooperate and which always defects as soon as you defect once) will never be able to establish *meta-cooperation*. Here, playing d has two significations: (i) refusing to cooperate; (ii) trying to *meta-cooperate*. To avoid this ambiguity must be the aim of all the strategies trying to reach a real success.
- It seems obvious that a good strategy must be able to accept only a first level cooperation if it cannot establish the second one.
- Reactivity like in the classical game seems necessary: you have to adjust yourself by taking in account rival's reactions.
- A good strategy should also be able to adjust itself to an opponent which plays c,d,c,d,c,d or c,c,d,d,c,c,d,d,c,c,d,d or any other scheme corresponding to an equitable reward's distribution. Taking into account all the meta-cooperations schemes seems to be very difficult and needs longer risk periods.
- In an ecological evolution, it is important to play as well as possible against oneself (this problem will be addressed in details).

About simplicity, graduality, memory, randomness, nothing seems "a priori" obvious.

4 Ecological competitions in homogeneous environment

In this section we search about strategies which obtains the best possible score in homogeneous environments, that is which are able to collect the best possible score when they play against themselves. Our conclusions are rather surprising.

Definition 1
We call *rational strategy* (resp.: *asymptotically rational strategy*) a strategy which when it plays against itself obtains the best possible score for every game length n (resp. asymptotically best possible score).

Formally we note $G(A;n)$ the score obtained by a strategy A when it plays against itself during n rounds (when the strategy is probabilistic $G(A;n)$ is the expectation of the A score on n rounds). By definition, a strategy is said to be *rational* if:

$$\forall n : G(A;n) = max\{G(X;n); X \text{ is a strategy}\}$$

By definition, a strategy is said *asymptotically rational* if:

$$lim_{n\to\infty}[G(A;n)/max\{G(X;n); X \text{ is a strategy}\}] = 1$$

A rational strategy in an ecological evolution with an homogeneous environment obtains the maximum possible reward. Such a strategy will have a score advantage when it will meet the others (especially in an ecological evolution starting from a sufficiently large amount of copies of itself).
To be efficient when you meet other strategies, you don't have to have the total score equally distributed between sisters strategies. This is why we will define the notions of *collectively rational* strategies and *asymptotically collectively rational* strategies.

Definition 2
We call *collectively rational* strategy (resp. *asymptotically collectively rational* strategy) a strategy which when two copies of itself plays together collects the maximum possible total score for every game length n (resp. asymptotically the best total score).

Formally we note $G'(A;n)$ the score obtained by two copies of the strategy A when they play together during n rounds (when the strategy is probabilist $G'(A;n)$ is the expectation of the score on n rounds).
By definition, a strategy is said to be *collectively rational* if:

$$\forall n : G'(A;n) = max\{G'(X;n); X \text{ is a strategy}\}$$

By definition, a strategy is said *asymptotically collectively rational* if:

$$lim_{n\to\infty}[G'(A;n)/max\{G'(X;n); X \text{ is a strategy}\}] = 1$$

4.1 The classical I.P.D. case

In the classical I.P.D. (parameters $S = 0$, $D = 1$, $C = 3$, $T = 5$) rational strategies are exactly those which never defect first (called nice strategies). Indeed to obtain the maximum possible score in a given panel each strategy must always play c (round [c,c]) which earns collectively 6 points each round (every change in this c sequence will reduce the collective score).
This is why, in ecological evolutions with a large variety of strategies only nice strategies stay alive (if all strategies have nearly the same number of representative, of course).
Asymptotically rational strategies are those which are able to obtain an average of 3 points by round when they play against themselves. They can defect sometimes, deterministically or probabilistically, but the defection number must tend to 0 to the infinity (for example, at the round n, defect with the $1/n$ probability).

In the classical I.P.D. the *collectively rational* (or *asymptotically collectively rational*) notion does not have any interest because to obtain the best possible score you need to play only [c,c] rounds which assure similar scores, thus *collectively rational* strategies are *rational* and *asymptotically rational* strategies are *collectively asymptotically rational*.
In the Lift Dilemma the situation is different because a strategy will have sometimes to sacrifice for the other. Some *collectively rational* strategies are not *rational*.

4.2 The Lift Dilemma case

In this section we present some mathematical results about the Lift Dilemma:

Theorem 1
If the two following equalities are satisfied: $S < P < R < T$ and $(S+T)/2 > R$, a *deterministic* strategy is never *rational* nor *asymptotically rational* nor *collectively rational* nor *asymptotically collectively rational*.

Proof.
When a *deterministic* strategy plays against itself there is never a round [c,d] or [d,c], thus it will always earn 3 points each round in the best case. We will see that there exist *probabilistic* strategies which earn an average of 4 points by round.

We will call *phased round* a round [c,c] or [d,d], and *unphased round* a round [c,d] or [d,c].

Theorem 2 (Rationality characterization)
A strategy is rational if and only if:

- at the first round and while the previous round is not unphased, it plays c with a $opti = 0,56696$ probability and d with $1 - opti$ probability;
- After the first unphased round, it uses a rule such that, for every possible past against itself, the one who played d at the first unphased round will play the opposite than the one who played c at the first unphased round (thus such a strategy when it plays against itself is deterministic after the first unphased round).

4.2.1 First examples

In this section periodic repetitions will be noted with a star. For example (c d d c d d c d d ...) will be noted (c d d)*.

Here is the simplest rational strategy called: REASON (the 0.56696 parameter is explained below).

REASON
• I play c with a 0,56696 probability and d with a 0,43304 probability at the first round and while the previous round is phased; • then if the first unphased round is [c d], I play (d c d c d c ...) denoted by (d c)* if the first unphased round is [d c], I play (c d c d c d ...) denoted by (c d)*

The following strategy called NAIVE-REASON is *collectively rational* but is not *rational*, thus an homogeneous NAIVE-REASON population will globally obtain the best possible score for an homogeneous population, but rewards will not be equally distributed between the entities.

NAIVE-REASON

- I play (c: 0,56696 ; d: 0,43304) at the first round and while the previous round is phased;
- then
  if the first unphased round is [c d], I play c*
  if the first unphased round is [d c], I play d*

This strategy can be explained by this way: "if at the first disagreement I have been exploited, I consider that I am a looser, and I accept to be always exploited. If at the first disagreement I win, I want to win every time".

Such a rule, used by individuals of the same species could be a mechanism able to create hierarchies. This kind of agreement respect the collective interest even if there is no equality between individuals. This strategy is more simple than the REASON strategy to obtain collective maximum score. Perhaps we could see here an explanation to the fact that democratical societies appeared more recently than despotic ones.

The following strategies try to improve REASON by being nicer. The idea is not to annoy easily offended strategies by trying first to cooperate.

GENTLE-REASON
I play like REASON excepted that the d probability during the first 3 rounds is equal to 0

REASON $[a, 1-a]$ (a parameter between 0 and 1)
I play like REASON excepted that I play c with a probability when I am waiting for an unphased round.

The two previous strategies are *asymptotically rational* because they loose only few points at the beginning of the game compared to what they can best expect. With a near 1 (for example 9/10) REASON$[a, 1-a]$, like GENTLE-REASON will avoid to annoy easily offended strategies.

While similar to REASON, the following strategy is neither *rational* nor *collectively rational* nor *asymptotically rational* nor *collectively asymptotically rational* because it satisfies itself too easily to obtain an average of 3 points each round. Of course we cannot expect this strategy to be a very good one.

COOP-REASON
I play COOPERATE (always c) until the first defection of my opponent, then start playing REASON.

4.2.2 Justification of the 0.56696 parameter

At first sight, the 0.56696 parameter seems strange. We explain here where does it come from. The 0.56696 number is the root of a polynomial equation obtained when trying to minimize the cost of *the period of search of an unphased round* (when the strategy plays randomly c or d) when a strategy plays against itself.
The fastest way to obtain an unphased round in this case is to play c and d with an equal probability of 1/2 (the average number of phased round is then 1).
But with this probability, we loose more points in average in each round than if we play (for example) with a 3/4 c probability and 1/4 d because [c,c] rounds which earn 3 points to each player are more frequent than [d,d] rounds which earn 1 point.
Thus it is not obvious that the best way to find an unphased round is to play c and d with a 1/2 probability. The mathematical study of this problem shows that the minimal unphased round searching cost is obtained for 0.56696 .
The computation of this cost leads to the following equation:

$$-(1-2p+2p^2)(p(R-P)+P-(T+S)/2)/[2p(1-p)]$$

which gives with our parameters $R = 3$, $P = 1$, $T = 8$, $S = 0$:

$$-(2p-3)(1-2p+2p^2)/2p(1-p)$$

We note that the gain obtained by 0.56696 instead of 0.5 is very small (less than 1/10 of a point). Thus for simplicity we can avoid it and make our searching period with a 1/2 probability.

4.2.3 Computation of the parameter 0.56966

Search for an unphased round when two strategies play c with probability p and d with probability $1-p$.

When we are in the period of search of an unphased round, [d,d] rands earns P points and the rands [c,c] earns R points. Thus during this period, in average we earn by round

$$Rp + P(1-p) = p(R-P) + P$$

Computation of the average length of this period:
length 0 [d,c] ou [c,d] probability $2p(1-p)$
length 1 ([d,d] ou [c,c]) followed by ([d,c] ou [c,d]) probability $2p(1-p)(p^2+(1-p)^2)$
length 2 ([d,d] ou [c,c]) two times followed by ([d,c] ou [c,d]) probability $2p(1-p)(p^2+(1-p)^2)^2$
etc.

Thus the expectation of length EL is:

$$EL = 2p(1-p)[(p^2+(1-p)^2) + 2(p^2+(1-p)^2)^2 + 3(p^2+(1-p)^2)^3 + +4(p^2+(1-p)^2)^4 + \ldots]$$

Comptation of the sum $X + 2X^2 + 3X^3 + 4X^4 + \ldots$

$$\begin{aligned} X + 2X^2 + 3X^3 + 4X^4 + \ldots &= X[1 + 2X + 3X^2 + 4X^3 + \ldots] \\ &= X[1 + X + X^2 + X^3 + \ldots]' \\ &= X[1/(1-X)]' \\ &= X/(1-X)^2 \end{aligned}$$

$$\begin{aligned} \text{Hence EL} &= 2p(1-p)(p^2+(1-p)^2)/(1-(p^2+(1-p)^2))^2 \\ &= 2p(1-p)(p^2+(1-p)^2)/(2p(1-p))^2 \\ &= (p^2+(1-p)^2)/(2p(1-p)) \\ &= (1-2p+2p^2)/2p(1-p) \end{aligned}$$

The loss L (compared to an immediate unphased round) is:

$$\begin{aligned} L &= [(1-2p+2p^2)/2p(1-p)][(T+S)/2 - (p(R-P)+P)] \\ &= (T+S)/2 - (p(R-P)+P)(1-2p+2p^2)/2p(1-p) \end{aligned}$$

With our parameters, we obtain:

$$L = (2p-3)(1-2p+2p^2)/2p(1-p)$$

For p=0.5 we find L=2 ; for p=0.7 we find L=2,209 ; for p=0.9 we find L=5,46 ; for p=0.4 we find L=2,38 ; for p=0.6 we find L=1,95 ; for p=0,55 we find L=1,938 ; for p=0,65 we find L=2,03

Minimum loss is obtained for $p = 0.5669640801\ldots$ The cheapest unphased round period searching is obtained when we play c with $p = 0.5669640801\ldots$ probability and d with $1-p$ probability.

4.2.4 Notes about determinism

We say in theorem 2 "for every possible past **against itself**" because it does not matter to play in opposite phase with other strategies when we are looking for *rational* strategies. It is only important for it to play in opposite phase **against itself**.

For example, consider the following strategy:

REASON-CAREFUL

- I play (c: 0.56696 ; d: 0.43304) at the first round and while the previous round is phased.
- Then
  - if the first unphased round is [c,d] I play (d c)* excepted if the opponent has played consecutively 3 d since the first unphased round; in this case I play d*
  - if the first unphased round is [d,c] I play (c d)* excepted if the opponent has played consecutively 3 d since the first unphased round; in this case I play d*

The excepted actions are never used when the strategy plays against itself and thus it is rational (theorem 2). The following strategy is also a rational one.

REASON-TIT-FOR-TAT

- I play (c: 0.56696 ; d: 0.43304) at the first round and while the previous round is phased;
- then I play TIT-FOR-TAT (I play what my opponent played on the previous move).

There is a generalization of the previous strategy. With a nearly 1 it will be less aggressive at the beginning of the game.

REASON-TIT-FOR-TAT $[a, 1-a]$

- I play (c: a ; d: $1-a$) at the first round and while the previous round is phased;
- then I play TIT-FOR-TAT

4.2.5 Remarks concerning the search of an unphased round

The $a = 0.56696$ parameter has been computed with an homogeneous population hypothesis (all the strategies are identical). Is this value always an optimal one in an heterogeneous population? The answer is no, we now explain why.

In an heterogeneous panel a new factor but be taken into account which favour a a value less than 0.5. It is clear that a strategy takes advantage to be the one which defects at the first unphased round, because if the rest of the game is during an odd number of rounds it will never get back the gain earned at the first unphased round. We can estimate to 4 points this kind of advantage (sometimes it will play an odd number of rounds thus it will earn 8 points, sometimes it will play an even number of rounds thus it will earn 0 points). This is really important in case of short games because the global earn is small.

Of course no optimal value of the parameter a can be mathematically determined because all depends of the starting panel. Simulations we have realized by confrontations between various REASON $[a, 1-a]$ have shown the following results:

- in an ecological evolution with 2 kinds of strategies: $a = 0.5$ wins against $a = 0.567$
- in an ecological competion with 5 kinds of strategies the order is: $a = 0.5$; $a = 0.433$; $a = 0.567$; $a = 0.1$; $a = 0.9$.

- in an ecological competion with 5 kinds of strategies the order is: $a = 0.567$; $a = 0.433$; $a = 0.5$; $a = 0.6$; $a = 0.7$; $a = 0.4$; $a = 0.3$; $a = 0.2$; $a = 0.1$; $a = 0.8$; $a = 0.9$

These results do not allow general conclusions excepted that choices of a near 0 or 1 are bad choices. Of course these results are very sensible to the random generator and in fact can change according to how experimentations are made (we have made averages on 1000 tournaments).

To obtain a good strategy you can of course watch the behaviour of the other player. For example, if your opponent has defected consecutively many times, it is certainly trying to exploite you. In such a case your insterest is to search again for a new unphased round. The following strategy is based on this idea with a test of 3 consecutive defections.

ITERATED-REASON

- I play (c: 0,56696 ; d: 0,43304) at the first round and while the previous round is phased.
- Then
  - if the first unphased round is [c,d], I play (d c)* excepted if my opponent has defected consecutively 3 times; in this case I forget the last phased round and I start again a search of unphased round,
  - if the first unphased round is [d,c], I play (c d)* excepted if my opponent has defected consecutively 3 times; in this case I forget the last phased round and I start again a search of unphased round.

4.2.6 Remarks about memory and complexity of strategies

The used of the first unphased round in the formulation of theorem 2 implies that rational strategies keep in memory the first unphased round. Thus rational strategies must not only be probabilistic but also must have an unlimited memory (they must remember the number and what they have played at the first unphased round).
To play well against oneself (that is to be rational) implies necessarily a certain level of complexity. To obtain a robust strategy, then the complexity of the strategy will be inscreased. Similar conclusions about the necessity of complexity have been already obtained concerning the classical I.P.D. [7, 11, 10]
The following experiments confirm the abstract analysis just described about the Lift Dilemma.

5 Practical study of confrontations

5.1 DEFECT against REASON

The REASON strategy plays well against itself, but it is not reactive (it does not take in account the behaviour of the opponent) after the first unphased round. Thus this strategy can be exploited for example by DEFECT strategy which always defect (t*). Let us show this result by hand:

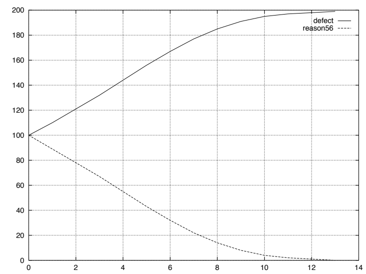
DEFECT against REASON gives:

[d,d] [d,d]... [d,d] [d,c] + [d,d] [d,c] [d,d] [d,c] [d,d] ...

After an unphased search period of REASON (before the +) which does not take a long time (one round in average), the confrontation continues with the REASON exploitation (REASON earns an average of 1 point each 2 rounds, while DEFECT earns 9 points each two rounds).

In a length game of 1000 rounds, globally, DEFECT against itself earns 1000 points, REASON against itself earns 4000, DEFECT against REASON earns 4500 while REASON earns 500.

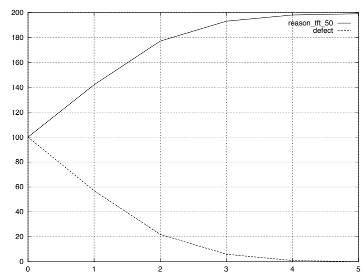
```
 Score list after tournament
1                defect =   5498
2              reason50 =   4498
```



5.2 DEFECT against REASON-TIT-FOR-TAT

The following experiment shows that the improvement added to REASON to obtain REASON-TIT-FOR-TAT leads to a strategy which now beats DEFECT.

```
 Score list after tournament
1   reason_tft_50 =   4997
2          defect =   2007
```



5.3 TIT-FOR-TAT against REASON

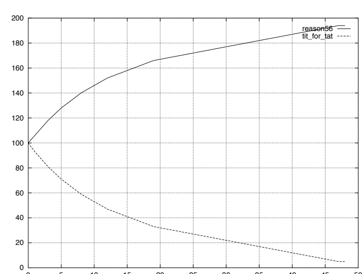
Once again [7] TIT-FOR-TAT reputation have to be reconsidered. In the Lift Dilemma, deterministic strategies cannot be good, and this the case for TIT-FOR-TAT. Nevertheless it is able to favour meta-cooperation with a period of two rounds. In fact it plays well against REASON (average of 4 points), unfortunately it is satisfied by 3 points against itself. In an ecological computation it is fatal for it.

TIT-FOR-TAT against RAISON gives:

[c,c] [c,c]... [c,c] [c,d] + [c,d] [d,c] [c,d] [d,c] [c,d] ...

In a length game of 1000 rounds, TIT-FOR-TAT against itseft earns 3000 points, REASON against itself earns 4000 points, TIT-FOR-TAT against REASON earns 4000 points like REASON.

```
 Score list after tournament
1        reason56 =   7996
2     tit_for_tat =   6997
```



5.4 REASON-TIT-FOR-TAT with TIT-FOR-TAT and DEFECT

REASON-TIT-FOR-TAT against TIT-FOR-TAT gives:

[c,c] [c,c]...[c,c] [c,d] + [d,c] [c,d] [d,c] ...

that is an average of 4 points each round.

REASON-TIT-FOR-TAT against DEFECT gives:

[d,d] [d,d] ... [d,d] [c,d] + [d,d] [d,d] [d,d] [d,d] ...

that is an average of 1 point each round together.

TIT-FOR-TAT against DEFECT:

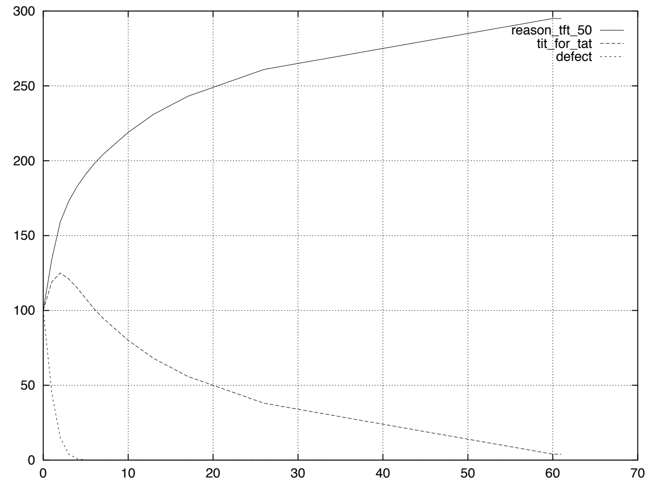
[c,d] [t,t] [t,t] [t,t] ...

that is 1 point each round together in average

DEFECT against itself earns 1 point each round
TIT-FOR-TAT against itself earns 3 points each round.
REASON-TIT-FOR-TAT against itself earns 4 points each round in average.

In an ecological confrontation between these 3 strategies, DEFECT disappears quickly, then TIT-FOR-TAT disappears.

```
 Score list after tournament
1   reason_tft_50 =   8997
2     tit_for_tat =   7996
3          defect =   3014
```



Note that TIT-FOR-TAT in the 5th first generations takes advantage of DEFECT's population decrease, but it can't do that a long time.

5.5 REASON-TIT-FOR-TAT against 10 basic strategies

To see if REASON-TIT-FOR-TAT is a good strategy, let us compare it in a tournament with 10 basic strategies. We first describe the 10 considered strategies.

COOPERATE always cooperates

DEFECT always defects

RANDOM cooperates with a probability of 0.5

TIT-FOR-TAT cooperates on the first move and then plays what its opponent played on the previous move

SPITE cooperates until the opponent defects, then defects all the time

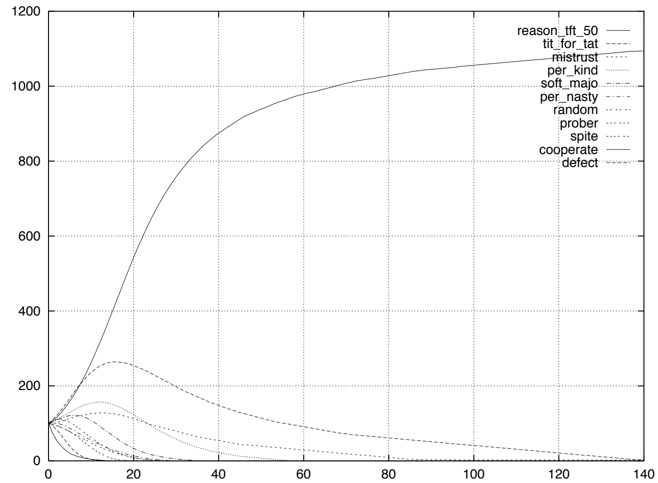
PER_KIND plays periodically [cooperate, cooperate, defect]

PER_NASTY plays periodically [defect, defect, cooperate]

SOFT_MAJO plays the opponent's most used move and cooperates in case of equality (first move considered as equality)

MISTRUST has the same behavior as TIT-FOR-TAT but defects on the first move

PROBER begins by playing [cooperate, defect, defect], then if the opponent cooperates on the second and the third move continues to defect, else plays TIT-FOR-TAT

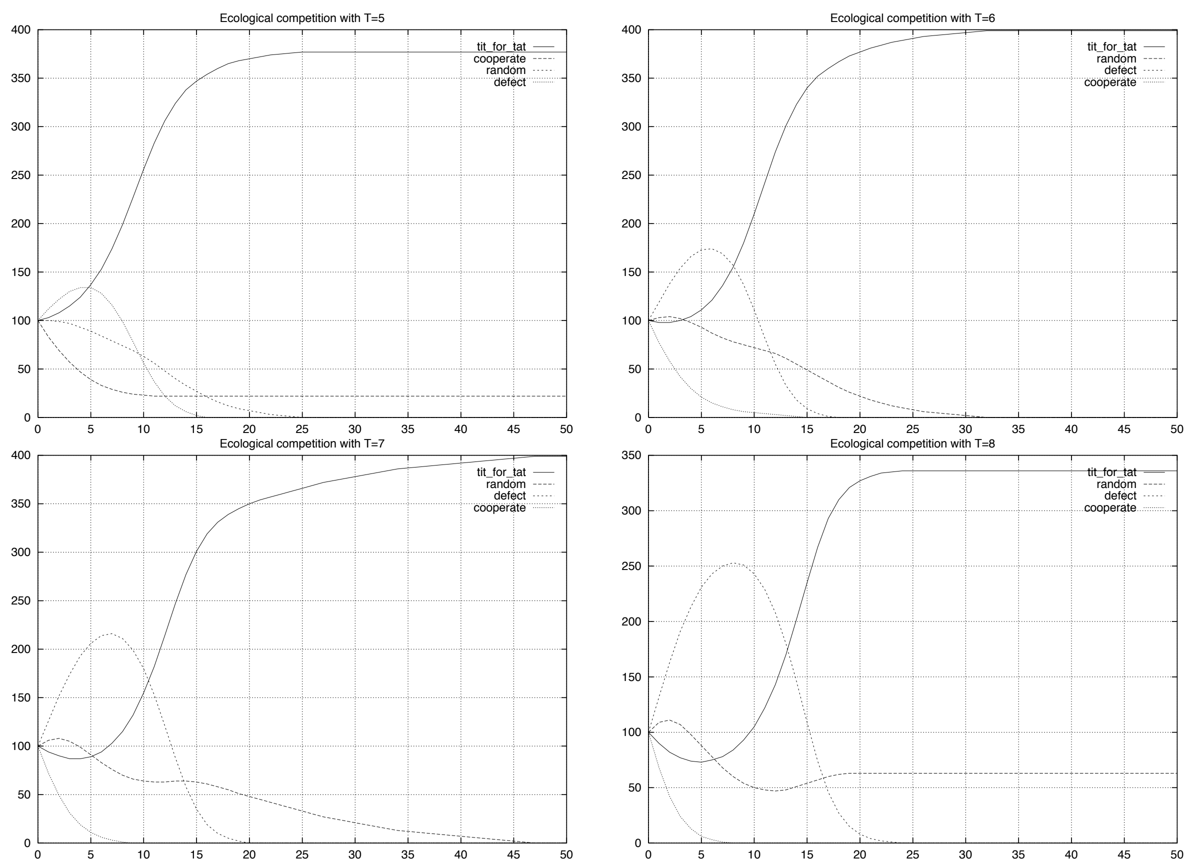


Note that during the first generations, TIT-FOR-TAT beats REASON-TIT-FOR-TAT, but once again, not for a long time.

6 Parameters sensibility

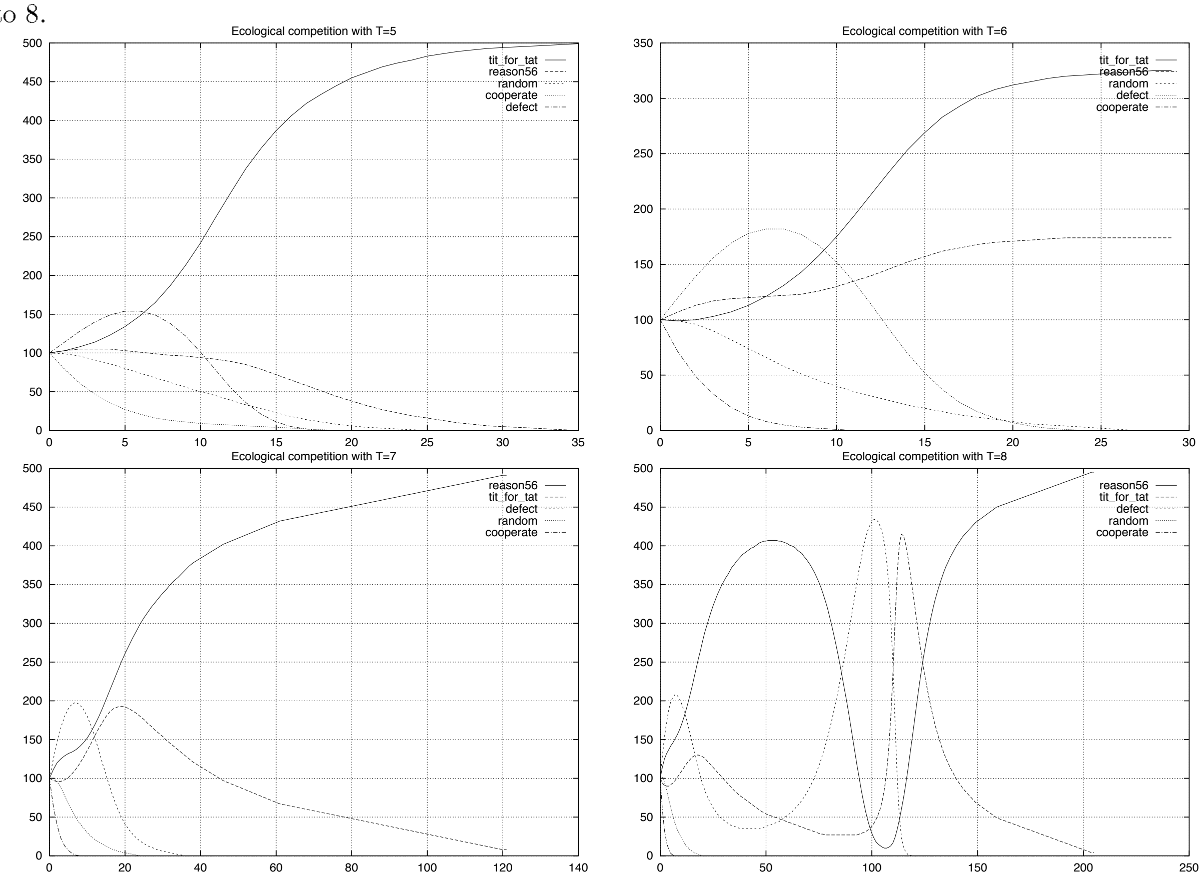
The 4 following graphics shows the influence of the T (tentation) parameter in ecological evolutions. To illustrate this influence let us first take 4 basic strategies without REASON and in a second time the same 4 strategies with REASON added to them.

In each case we change the T parameter from $T = 5$ (classical I.P.D.) to $T = 6$, $T = 7$ and finally $T = 8$ (Lift Dilemma).



We can see that to be probabilistic is not sufficient to be good in the cases where $T > (R + D)/2$ (see RANDOM).

Now we add REASON to the previous panel, and we change the T parameter once again from 5 to 8.



6.1 Conclusion

In this paper we have shown that a very small change in Iterated Prisoner's Dilemma payoff matrix leads to an iterated game which properties are very different than those of the classical IPD. Two levels of cooperation are possible in this game. This creates an iterated game much more difficult to analyse than the classical I.P.D. Nevertheless very concrete situations of social life are simulated with it. One of our conclusions mathematically proved is that only probabilistic strategies can make a hight score when the play against themselves. We have then found interesting characteristics allowing us to define good strategies like REASON or REASON-TIT-FOR-TAT. Building good strategies for the Lift Dilemma is now much more interesting and complex than for the classical game.
A simulation software with many strategies is already available for Unix, Dos or Windows by web at `http://www.lifl.fr/~mathieu/ipd` or by anonymous ftp on the following site `ftp.lifl.fr` in `pub/users/mathieu/soft`

7 Bibliography

References

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

[2] R. Axelrod. The evolution of strategies in the iterated prisoner's dilemma. In L. Davis, editor, *Genetic Algorithms and the Simulated Annealing*, chapter 3, pages 32–41. Pitman, London, 1987.

[3] R. Axelrod and D. Dion. The further evolution of cooperation. *Science*, 242:1385–1390, 1988.

[4] R. Axelrod and W. D. Hamilton. The evolution of cooperation. *Science*, 211:1390–1396, 1981.

[5] S. Bankes. Exploring the foundations of artificial societies. In R. A. Brooks and P. Maes, editors, *Artificial Life, Proc. 4th International Workshop on the Synthesis and Simulation of Living Systems*, volume 4, pages 337–342. MIT Press, 1994.

[6] J. Batali and P. Kitcher. Evolutionary dynamics of altruistic behavior in optional and compulsory versions of the iterated prisoner's dilemma. In R. A. Brooks and P. Maes, editors, *Artificial Life, Proc. 4th International Workshop on the Synthesis and Simulation of Living Systems*, volume 4, pages 344–348. MIT Press, 1994.

[7] B. Beaufils, J.P. Delahaye, and P. Mathieu. Our meeting with gradual, a good strategy for the iterated prisoner's dilemma. In C. Langton, editor, *Artificial Life, Proc. 5th International Workshop on the Synthesis and Simulation of Living Systems*, volume 5, 1996. Artificial Life 5, Nara, Japan, May 16-18 (proceedings to appear).

[8] J. Bendor. In good times and bad: Reciprocity in an uncertain world. *American J. of Political Science*, 31:531–558, 1987.

[9] Robert Boyd and Jeffrey P. Lorberbaum. No pure strategy is evolutionarily stable in the repeated prisoner's dilemma game. *Nature*, 327:58–59, 1987.

[10] J.P. Delahaye and P. Mathieu. L'altruisme perfectionné. *Pour La Science (French edition of Scientific American)*, 187:102–107, 1993.

[11] J.P. Delahaye and P. Mathieu. Complex strategies in the iterated prisoner's dilemma. In A. Albert, editor, *Chaos and Society*, Amsterdam, 1995. IOS Press.

[12] M. R. Frean. The prisoner's dilemma without synchrony. *Proc. Royal Society London*, 257(B):75–79, 1994.

[13] H. C. J. Godfray. The evolution of forgiveness. *Nature*, 355:206–207, 1992.

[14] N. V. Joshi. Evolution of cooperation by reciprocation within structured demes. *J. of Genetics*, 66(1):69–84, 1987.

[15] K. Lindgren. Evolutionary phenomena in simple dynamics. In J.D. Farmer C. Langton, C. Taylor and S. Rasmissen, editors, *Proc. Workshop on Art. Life*, volume 2, pages 295–312. Santa Fe Institute, Addisson Wesley, 1992.

[16] C. Martino. Emergent nastiness in iterated prisoner's dilemma games. 2.725: Design and Automation, 1995.

[17] R. M. May. More evolution of cooperation. *Nature*, 327:15–17, 1987.

[18] P. Molander. The optimal level of generosity in a selfish, uncertain environment. *J. of Conflict Resolution*, 29(4):611–618, 1985.

[19] M. Nowak. Stochastic strategies in the prisoner's dilemma. *Theoretical Population Biology*, 38:93–112, 1990.

[20] M. Nowak and K. Sigmund. The evolution of stochastic strategies in the prisoner's dilemma. *Acta ApplicandæMathematicæ*, 20:247–265, 1990.

[21] M. Nowak and K. Sigmund. Tit for tat in heterogeneous populations. *Nature*, 355:250–253, 1992.

[22] M. Nowak and K. Sigmund. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature*, 364:56–58, 1993.

[23] R. Pool. Putting game theory to the test. *Science*, 267:1591–1593, 1995.

[24] M. D. Smucker, E. Ann Stanley, and D. Ashlock. Analyzing social network structures in the iterated prisoner's dilemma with choice and refusal. RR. CS-TR-94-1259, University of Wisconsin-Madison, Departement of Computer-Sciences, 1994.

[25] Xin Yao and P. J. Darwen. An experimental study of n-person iterated prisoner's dilemma game. *Informatica*, 18:435–450, 1994.