CHAPTER 11

The Iterated Lift Dilemma

How to Establish Meta-Cooperation with your Opponent

J.P. Delahaye, P. Mathieu, and B. Beaufils

Laboratoire d'Informatique Fondamentale de Lille

A very small change in the Iterated Prisoner's Dilemma (IPD) payoff matrix leads to an iterated game called the Iterated Lift Dilemma[1] the properties of which are very different from those of the classical IPD (CIPD). We show that the following ideas 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 high score when they play against themselves; (iii) complex dynamics can appear (at the *edge of chaos*) as soon as three strategies are confronted. Our idea, already argumented in the case of the CIPD, 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 CIPD.

1. Introduction

Conflicting situations are not only a driving force in nature and society, they are also the entry points for many investigations in Artificial Intelligence especially in Distributed Artificial Intelligence, Multi Agents Systems, formal model of rational action, CSCW, concurrent engineering and HCI.

The Iterated Prisoner's Dilemma is a model for studying cooperation and conflicts. It's an iterated game.

An *iterated game* is a game with two players A and B (also called *strategies*) who play an unknown finite number of rounds. On each round, each player chooses between two actions `C` (for *Cooperation*), and `D` (for *Defection*). 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 `C` and the player B plays `D` is noted [`C`,`D`]; a round where the player A plays `D` and the player B plays `C` is noted [`D`,`C`]; and finally a round where both players play `D` is noted [`D`,`D`].

When the players play the round number n they play simultaneously taking into account the game history (that is all the preceding choices they both have made at all rounds i with $i < n$).

[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*

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 long enough (> 10) to allow interesting phenomena and to obtain robust results.

The results of the player's choices are quantified by a payoff matrix, which is shown in table 1.1. The parameter R is the reward for mutual cooperation (round [`C`,`C`]), T is the temptation to defect against an opponent cooperation, who then gets the sucker's payoff S (round [`D`,`C`]). In case of mutual defection both get the punishment P (round [`D`,`D`]).

As players can now choose a way of playing before the game begins, or during the game, they can be said having a *strategy*.

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), the points obtained by each player are cumuled. The winner is the player who has the greatest score.
- There are *round-robin tournaments*. We take k strategies, each one playing against all the others (including itself) in single confrontations. Points of confrontations are cumuled, the winner is the player who has the greatest score.
- There are *ecological evolutions*. For example we consider only 3 strategies `A`, `B`, `C`; we start with a population of one hundred players choosing the `A` strategies, one hundred playing `B` strategies and one hundred playing `C` strategies, that is 300 entities (this defines the first generation); a round-robin tournament is computed (as if each entity plays against the 299 other entities); scores are computed for each strategy (sum of all scores from entities of the same kind); a new population is then computed for each strategy which is proportional to the score obtained. To simplify we consider that the total population is constant (here 300). This defines the second generation. This computation is repeated until populations become stable.

We could let the total population 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*, nice[2] strategies proliferate and replace bad ones. To be a good strategy in an *ecological evolution* a strategy needs 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 evolution* are really robust ones.

In the CIPD the following parameters values are generally used:

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

[2] *nice* strategies, as opposed to *bad* or *naughty* ones, are the strategies which never defect prior to its opponents.

Table 1.1. Iterated Lift Dilemma payoff matrix. Row player scores are given first.

	Cooperate	Defect
Cooperate	$R = 3$, $R = 3$ *Reward* for mutual cooperation	$S = 0$, $T = 8$ *Sucker*'s payoff *Temptation* to defect
Defect	$T = 8$, $S = 0$ *Temptation* to defect *Sucker*'s payoff	$P = 1$, $P = 1$ *Punishment* for mutual defection

which obey to the next two fundamental inequations:

$$S < P < R < T$$

and

$$S + T < 2R$$

The first one says that the one shot game is a dilemma, whereas the second is used to favor cooperation in the iterated version.

This game has been found to be a very good way of studying cooperation and evolution of cooperation. A theory of cooperation based upon reciprocity has been set in a wide literature, such as in (Axelrod 1984; Axelrod and Dion 1988; Axelrod and Hamilton 1981).

Experimental studies of the IPD and its strategies need a lot of computation time. Thanks to the progress of computer science and computers, a lot of scientists have studied it as they have been able to use specific methods, like genetic and evolutionary algorithms, see (Axelrod 1987; Bankes 1994; Boyd and Loberbaum 1987; Lindgren 1992; Martino 1995; Nowak and Sigmund 1992; Nowak and Sigmund 1993; Smucker, Stanley, and Ashlock 1994; Yao and Darwen 1994).
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: (Batali and Kitcher 1994; Bendor 1987; Frean 1994; Godfray 1992; Joshi 1987; May 1987; Molander 1985; Nowak and Sigmund 1990; Pool 1995; Nowak 1990).

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

$$S + T > 2R$$

For instance we will study the following parameters, which are shown on table 1.1:

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

This small change on the classical game entails many surprising consequences. It becomes now more interesting to agree with its rival for playing [`C`,`D`] then

[D,C] then [C,D] then [D,C] etc. (which gives an average payoff of $(S + T)/2$, 4 points each round for each player) than playing [C,C] then [C,C] then [C,C] etc. (which only gives an average payoff of R, 3 points each round for each player).

It is still a dilemma: due to the first classical inequality that has not been changed, the collective interest contradicts the individual one. To maximize reward needs a subtle agreement.

As for the CIPD 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 `all_d` (which always defects) never looses against any other strategy but scores very few points each time (it wins against its opponents but each game its moves 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 T, 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 R (3) points by round for each player;
– the upper level (or *meta-cooperation* level): to find a way of agreeing with the opponent to win and loose alternatively, that is to play [C,D] [D,C] then [C,D] then [D,C] etc., which gives an average reward of $(T + S)/2$ (4) points by round for each player.

To have success in meta-cooperation each player must play in *opposite phase* [C,D] then [D,C] then [C,D] then [D,C] etc. which is difficult because it needs some *coordination* and a great risk of loss for the player who plays C first (it could wait reciprocity for a long time !)

Other high-level cooperations are also possible: for example, playing [C,D] [C,D] [D,C] [D,C] [C,D] [C,D] [D,C] [D,C] etc. (periodicity 4). Such meta-cooperations need 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 any kind of 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 (Frean 1994).

The *lift dilemma* does not take into 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 CIPD.

Numerous situations with humans or artificial agents can be represented by this game. In particular, every situation where an object is periodically given to the two players, with the possibility that one (and only one) of them takes it, or that no one takes it.

Even if it seems that this new dilemma is very similar to the CIPD, we show that this is not the case and that this game has new surprising properties.

2. Real examples of Iterated Lift Dilemma

Here are some 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. The chances of each candidate to be elected are evaluated to 30%.
- [D,D]. The two candidates take place in an aggressive competition which damages their reputations. Their fight scares some electors and the party globally looses votes. Now 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. 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 up to 5 or 10.

2.2 Collaborator's recruitment session

During collaborator's recruitment session (in research centers, universities, or private companies) 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 will help you to support yours. Next time you will support my future candidate*".

A round [C,C] is a session where each team kindly defends its candidate, a round [D,D] is a session where the two teams roughly fight for the recruitment of its candidate (with a great risk to see the candidate of a third team be chosen), a round [D,C] is a session where one of the two teams leaves its chances to 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 to buy alternatively the objects, instead of out-bidding mutually which leads to 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 save our money*".

A round [C,C] is an auction where they try quietly to buy the art object. A round [D,D] is an auction where they wildly out-bid to have the object. A round [D,C] is an auction where one of the two buyers abstain from saying anything hoping a feedback next time.

2.4 The two music amateur neighbors

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

- [C,D] I can ear my music and I am not disturbed by yours. I obtain a satisfaction of 8 "pleasure points".
- [D,C] You can ear your music and I can't ear mine. It counts for 0 "pleasure points" for me.
- [D,D] I am trying to ear my music but I can simultaneously ear yours. I just obtain 1 "pleasure point" but you too !
- [C,C] We both renounce to listen to 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 avoids 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 specy, 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 (for example between two males which desire the same female). A round [`D`,`D`] corresponds to a violent fight which can result in severe wounds.

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 wins 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 a good strategy be ?

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

– As for the CIPD, 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 be successful the players must be able to establish cooperation (or meta-cooperation) to obtain the maximum global score.
– It is a real dilemma: in the Game Theory meaning, 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 `all_d` (which always defects). By construction, the game is more difficult than the CIPD due to the existence of two levels of cooperation and also to the high 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 scare an anxious opponent which will think that you don't want to cooperate. For example the `spiteful` strategy (which cooperates while 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 of cooperation if it cannot establish the second one.
- Reactivity, as in the classical game, seems necessary: you have to adjust yourself by taking into account the rival's reactions.
- A good strategy should also be able to adjust itself to an opponent which plays (`CDCDCD`) or (`CCDDCCDDCCDD`) or any other scheme corresponding to an equitable reward's distribution. Taking into account all meta-cooperation 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 later).

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

4. Ecological evolutions in homogeneous environment

In this section we look for strategies which obtain 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 4.1. *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).*

Let us note $\mathrm{V}_n(\mathrm{A})$ the score obtained by a strategy `A` when it plays against itself during n rounds (when the strategy is probabilistic $\mathrm{V}_n(\mathrm{A})$ is the expectation of the `A` score on n rounds).

By definition, a strategy is said to be *rational* if:

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

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

$$\lim_{n\to\infty}\left[\frac{V_n(A)}{\max\{V_n(X); X \text{ is a strategy}\}}\right] = 1$$

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, the total score does not have to be equally distributed between sister strategies. This is why we will define the notions of *collectively rational* strategies and *asymptotically collectively rational* strategies.

Definition 4.2. *We call* collectively rational *strategy (resp.* asymptotically collectively rational *strategy) a strategy which collects the maximum possible total score for every game length n (resp. asymptotically the best total score) when two copies of itself play together.*

Formally we note V'$_n$(A) the score obtained by two copies of the strategy A when they play together during n rounds (when the strategy is probabilist V'$_n$(A) is the expectation of the score on n rounds).

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

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

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

$$\lim_{n\to\infty}\left[\frac{V'_n(A)}{\max\{V'_n(X); X \text{ is a strategy}\}}\right] = 1$$

4.1 The CIPD case

In the CIPD (parameters $S = 0$, $P = 1$, $C = 3$, $T = 5$) rational strategies are exactly those which never defect first (called nice strategies). Indeed to obtain the maximum possible score in a given set each strategy must always play `C` (round [`C`,`C`]) which scores collectively $2R$ (6) points each round (every change in this `C` sequence will reduce the collective score).

That explains that, in ecological evolutions involving a large variety of strategies, only nice strategies stay alive (if all strategies have nearly the same number of representatives, of course).

Asymptotically rational strategies are those which are able to obtain an average of R (3) points by round when they play against themselves. They can defect sometimes, deterministically or probabilistically, but the defection number must tend from the infinity to 0 (for example, at the round n, defect with the $1/n$ probability).

In the CIPD 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 4.1. *If the two following equalities are satisfied: $S < P < R < T$ and $S+T > 2R$, 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 score R (3) points each round in the best case. We will see that there exist *probabilistic* strategies which score 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 4.2. *(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 history 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).*

Proof. The former point will be established later with the justification and computation of the 0.56696 probability, while the latter one is clearly obvious.

4.2.1 First examples. In this section periodic repetitions will be noted with a star. Let us note, for example, the moves (CDDCDDCDD)... as (CDD)$*$.

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 (DC)$*$
  – if the first unphased round is [D,C], I play (CD)$*$

The following strategy called `naive-reason` is *collectively rational* but is not *rational*. An homogeneous `naive-reason` population will globally obtain the best possible score for an homogeneous population. Rewards will however 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 always be 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 respects 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 is a parameter between 0 and 1)
I play like `reason` excepted that I play C with probability a when I am waiting for an unphased round.

The two previous strategies are *asymptotically rational* because they only lose 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 `all_c` (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 it comes from. This 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 lose more points on average in each round than if we play (for example) `C` with a 3/4 probability and `D` with one of 1/4 `D` because [`C`,`C`] rounds which give 3 points to each player are more frequent than [`D`,`D`] rounds which pay only 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 0.56696. 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`] rounds score P points and the rounds [`C`,`C`] score R points. Thus during this period, on average we score by round

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

Computation of the average length of this period:

length	moves	probability
0	[`D`,`C`] or [`C`,`D`]	$2p(1-p)$
1	([`D`,`D`] or [`C`,`C`]) followed by ([`D`,`C`] or [`C`,`D`])	$2p(1-p)(p^2+(1-p)^2)$
2	([`D`,`D`] or [`C`,`C`]) two times followed by ([`D`,`C`] or [`C`,`D`])	$2p(1-p)(p^2+(1-p)^2)^2$
$\vdots$	$\vdots$	$\vdots$

Thus the expectation of length EL is:

$$\begin{aligned} \text{EL} &= 2p(1-p)\left[(p^2+(1-p)^2) + 2(p^2+(1-p)^2)^2\right. \\ &\quad \left. + 3(p^2+(1-p)^2)^3 + 4(p^2+(1-p)^2)^4 + \cdots\right] \end{aligned}$$

Computation of the sum $X + 2X^2 + 3X^3 + 4X^4 + \cdots$

$$\begin{aligned} X + 2X^2 + 3X^3 + 4X^4 + \cdots &= \\ &= X[1 + 2X + 3X^2 + 4X^3 + \cdots] \\ &= X[1 + X + X^2 + X^3 + \cdots]' \\ &= X\left[\frac{1}{(1-X)}\right]' \\ &= \frac{X}{(1-X)^2} \end{aligned}$$

$$\begin{aligned} \text{Hence, EL} &= \frac{2p(1-p)\left(p^2+(1-p)^2\right)}{\left[1-(p^2+(1-p)^2)\right]^2} \\ &= \frac{2p(1-p)\left(p^2+(1-p)^2\right)}{\left(2p(1-p)\right)^2} \\ &= \frac{\left(p^2+(1-p)^2\right)}{(2p\,(1-p))} \\ &= \frac{1-2p+2p^2}{2p(1-p)} \end{aligned}$$

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

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

With our parameters, we obtain:

$$L = \frac{(2p-3)(1-2p+2p^2)}{2p(1-p)}$$

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 4.2 "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 (`DC`)$*$ 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 (`CD`)$*$ 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-[a,1-a]-tit_for_tat`

- 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 which favors a value less than 0.5 for a must be taken into account. It is in fact clear that a strategy takes advantage to be the one which defects at the first unphased round, because if the rest of the game lasts an odd number of rounds it will never get back the gain scored at the first unphased round. We can estimate to 4 points this kind of advantage (sometimes it will be played an odd number of rounds thus it will score 8 points, sometimes it will last an even number of rounds thus it will earn 0 points). This is really important in case of short games because the global payoff is small.

Of course no optimal value of the parameter a can be mathematically determined because everything depends of the starting panel. The 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 competition 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 competition 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 except 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 an average of 1000 tournaments).

To obtain a good strategy you can of course watch the behavior of the other player. For example, if your opponent has consecutively defected many times, it is certainly trying to exploit you. In such a case your interest 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 (`DC`)$*$ 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 (`CD`)$*$ 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 use 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 must also have a 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.

Then to obtain a robust strategy, the complexity of the strategy will be increased. Similar conclusions about the necessity of complexity have already been obtained concerning the CIPD (Beaufils, Delahaye, and Mathieu 1996; Delahaye and Mathieu 1995; Delahaye and Mathieu 1993).

The following experiments confirm the abstract analysis just described about the Lift Dilemma.

5. Practical study of confrontations

5.1 `all_d` against `reason`

The `reason` strategy plays well against itself, but is not reactive (it does not take into account the behavior of the opponent) after the first unphased round. Thus this strategy can be exploited for example by `all_d` strategy which always defect (D)∗. Let us show this result.

`all_d` 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 on average), the confrontation continues with the `reason` exploitation (`reason` scores an average of 1 point each 2 rounds, while `all_d` scores 9 points each two rounds).

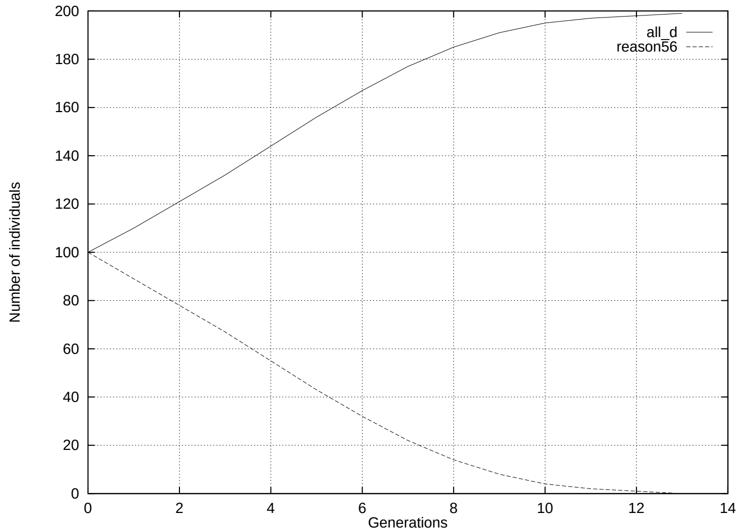
In a length game of 1000 rounds, globally, `all_d` against itself scores 1000 points, `reason` against itself scores 4000, `all_d` against `reason` scores 4500 while `reason` scores 500. See figure 5.1.

5.2 `all_d` 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 `all_d`. See figure 5.2.

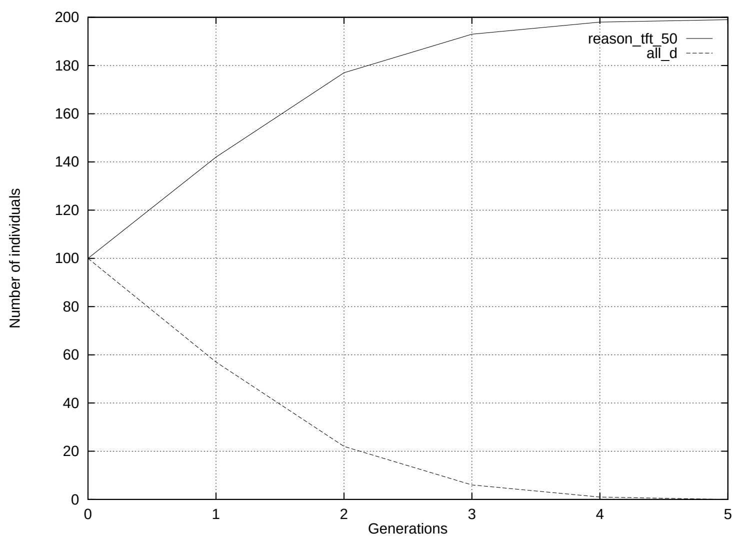
5.3 `tit_for_tat` against `reason`

Once again, see (Beaufils, Delahaye, and Mathieu 1996) for previous questions, `tit_for_tat` reputation have to be reconsidered. In the Lift Dilemma, deterministic strategies cannot be good, and this is the case for `tit_for_tat`. Nevertheless it is able to favor 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.




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

Fig. 5.1. `all_d` vs. `reason`

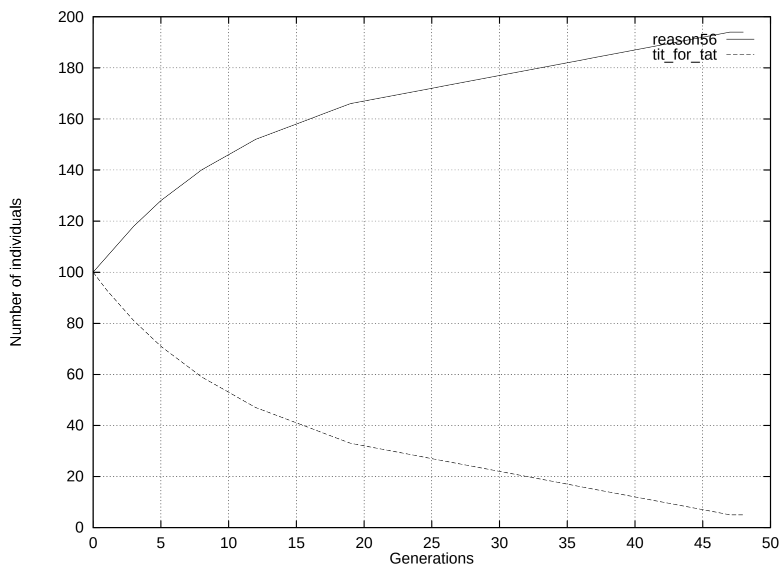



```
Score list after tournament
1      reason_tft_50 =   4997
2              all_d =   2007
```

Fig. 5.2. `all_d` vs. `reason-tit_for_tat`

tit_for_tat against reason 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 itself scores 3000 points, reason against itself scores 4000 points, tit_for_tat against reason scores 4000 points like reason. See figure 5.3.




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

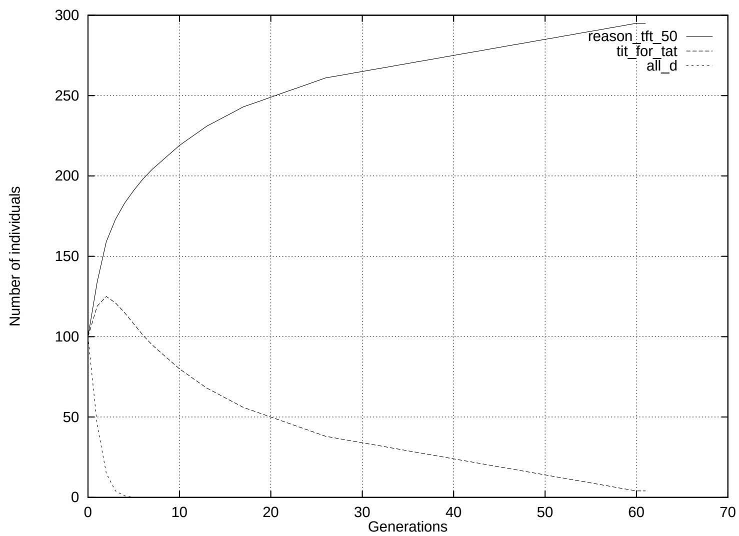
Fig. 5.3. tit_for_tat vs. reason

5.4 reason-tit_for_tat with tit_for_tat and all_d

- 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 all_d 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 all_d gives [C,D][D,D][D,D][D,D]··· that is 1 point each round together in average
- all_d against itself scores 1 point each round
- tit_for_tat against itself scores 3 points each round

– `reason-tit_for_tat` against itself scores 4 points each round in average.

In an ecological evolution involving these 3 strategies, `all_d` disappears quickly, then `tit_for_tat` disappears.




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

Fig. 5.4. `reason-tit_for_tat`, `tit_for_tat`, and `all_d`

Note, as shown on figure 5.4, that `tit_for_tat` until the 5th first generations takes advantage of `all_d`'s population decrease, but it can't do that a long time.

5.5 `reason-tit_for_tat` against 10 basic strategies

In order to determine if `reason-tit_for_tat` is a good strategy, let us compare it in an ecological evolution with 10 basic strategies. We first describe the 10 considered strategies:

`all_c` always cooperates
`all_d` always defects
`ipd_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
`spiteful` cooperates until the opponent defects, then defects all the time

per_ccd plays periodically [cooperate, cooperate, defect]

per_ddc 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

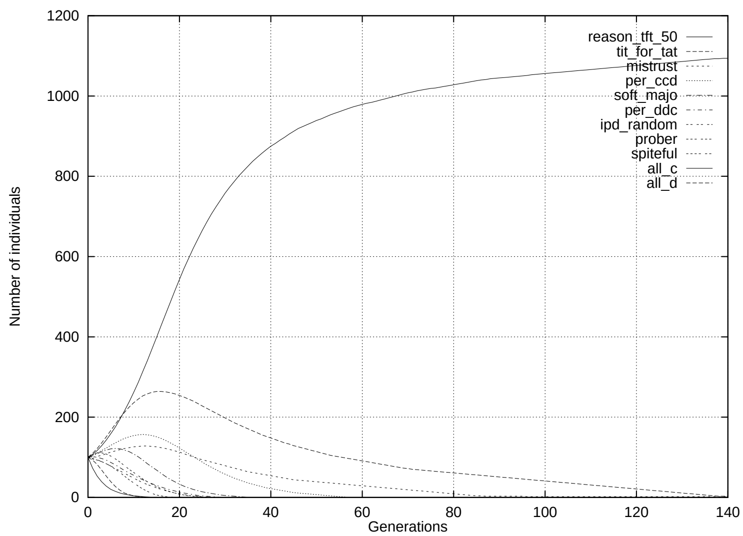


Fig. 5.5. reason-tit_for_tat against 10 basic strategies

As shown on the figure 5.5, during first generations, tit_for_tat beats reason-tit_for_tat, but once again, not for a long time.

6. Parameters sensibility

Graphics on figure 6.1 shows the influence of the T (temptation) parameter in ecological evolutions.

To illustrate this influence let us, in a first time, 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 $T = 5$ (CIPD) to $T = 6$, $T = 7$ and finally $T = 8$ (ILD).

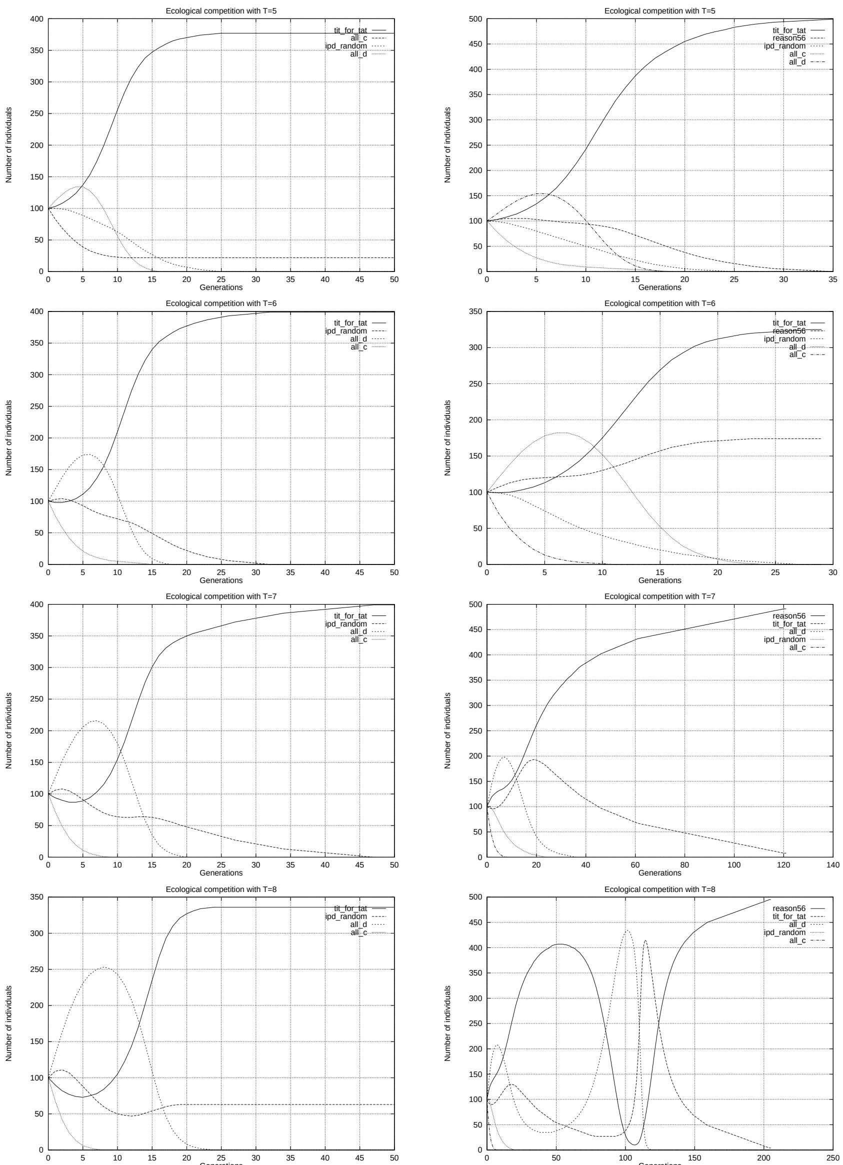


Fig. 6.1. Sensibility to the T parameter

We can see that to be probabilistic is not sufficient to be good in the cases where $2R < S + T$ (see `ipd_random`).

It is also to be quoted that such phenomena could also be seen with the classical Iterated Prisoner's Dilemma.

7. Conclusion

In this chapter 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 CIPD. Two levels of cooperation are possible in this game. This creates an iterated game much more difficult to analyze than the classical IPD 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 high score when they 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.

The simulation software we use for the experiments, is already available, with many strategies, for Unix, DOS or Windows computer system architectures on the World Wide Web at `http://www.lifl.fr/IPD` or by anonymous ftp on the following site `ftp.lifl.fr` in `pub/projects/IPD`.

References

Axelrod, R. (1984). *The Evolution of Cooperation.* New York, USA: Basic Books.

Axelrod, R. (1987). The Evolution of Strategies in the Iterated Prisoner's Dilemma. In L. Davis (Ed.), *Genetic Algorithms and the Simulated Annealing*, Chapter 3, pp. 32–41. London, UK: Pitman.

Axelrod, R. and D. Dion (1988). The Further Evolution of Cooperation. *Science 242*, 1385–1390.

Axelrod, R. and W. D. Hamilton (1981). The Evolution of Cooperation. *Science 211*, 1390–1396.

Bankes, S. (1994). Exploring the Foundations of Artificial Societies. See Brooks and Maes (1994), pp. 337–342. Artificial Life 4, Cambridge, MA, USA, July 6-8 1994.

Batali, J. and P. Kitcher (1994). Evolutionary Dynamics of Altruistic Behavior in Optional and Compulsory Versions of the Iterated Prisoner's Dilemma. See Brooks and Maes (1994), pp. 344–348. Artificial Life 4, Cambridge, MA, USA, July 6-8 1994.

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

Bendor, J. (1987). In Good Times and Bad: Reciprocity in an Uncertain World. *American J. of Political Science 31*, 531–558.

Boyd, R. and J. P. Loberbaum (1987). No Pure Strategy is Evolutionarily Stable in the Repeated Prisoner's Dilemma Game. *Nature 327*, 58–59.

Brooks, R. A. and P. Maes (Eds.) (1994). *Artificial Life IV: Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems*. Cambridge, MA, USA: A Bradford Book The MIT Press. Artificial Life 4, Cambridge, MA, USA, July 6-8 1994.

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

Delahaye, J. and P. Mathieu (1995). Complex Strategies in the Iterated Prisoner's Dilemma. In A. Albert (Ed.), *Chaos and Society*, Volume 29 of *Frontiers in Artificial Intelligence and Applications*, Amsterdam, Netherlands, pp. 283–292. Universit du Qubec Hull, Canada: IOS Press/Presses de l'Universit du Qubec. Chaos & Society 1994, Trois Rivires, Canada, June 1-2 1994.

Frean, M. R. (1994). The Prisoner's Dilemma without Synchrony. *Proc. Royal Society London 257*(B), 75–79.

Godfray, H. C. J. (1992). The Evolution of Forgiveness. *Nature 355*, 206–207.

Joshi, N. V. (1987). Evolution of Cooperation by Reciprocation within Structured Demes. *Journal of Genetics 66*(1), 69–84.

Lindgren, K. (1992). Evolutionary Phenomena in Simple Dynamics. In C. G. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen (Eds.), *Artificial Life II: Proceedings of the Second Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems*, Volume 10 of *Santa Fe Institute Studies in the Sciences of Complexity*, Reading, MA, USA, pp. 295–312. Addisson–Wesley Publishing Company. Artificial Life 2, Santa Fe, USA, February 1990.

Martino, C. (1995). Emergent Nastiness in Iterated Prisoner's Dilemma Games. 2.725: Design and Automation.

May, R. M. (1987). More Evolution of Cooperation. *Nature 327*, 15–17.