Influence of Reputation on Resistance against Bad Agents in the Iterated Prisoner's Dilemma

Per Knops, **Neil Yorke-Smith**
TU Delft

Abstract

In the Iterated Prisoner's Dilemma players can take advantage of other players. This has no drawbacks for the player after that game, since it is assumed that the players have no memory. When reputation is introduced however, a single game of the Prisoner's Dilemma can influence other games. In this paper research is done on how reputation can influence the population of a spatial Iterated Prisoner's Dilemma. This is then extended by adding groups. While in total the amount of 'bad players' decreases and the amount of 'good players' increases, the cooperation does not always increase.

1 Introduction

The Iterated Prisoner's Dilemma (IPD) is an iterative two-player-game. This dilemma describes a trade-off between cooperation and defection over a certain number of rounds. Every round each player chooses either to cooperate or to defect. In the case that both players decide to cooperate, they both receive a reward R. If both players decide to defect, they both receive a smaller reward P (punishment payoff). If only one player decides to cooperate while the other player defects, that player will receive S, also called the sucker's payoff, but the other player will receive a bigger reward T (temptation payoff). These entries must satisfy the following relationships:

$$T > R > P > S$$

and

$$2R > T + S$$

	B cooperates	B defects
A cooperates	R (A), R (B)	S (A), T (B)
A defects	T (A), S (B)	P (A), P (B)

Table 1: Payoff matrix for a two-player IPD

If a single round of the Prisoner's Dilemma game is played, mutual defection is the only Nash-equilibrium. In the case of an Iterated Prisoner's Dilemma however this is not always the case. If the number of rounds played is known and very small, always defecting will still probably be the strategy of both players. When the number of rounds played is not very small, cooperation will improve the overall score, since opponents will be more likely to cooperate later on as well.

Two extensions of this IPD will be examined in this paper: A spatial configuration and reputation. Both extensions change the way the IPD works, and will therefore change the way the players play the game.

The first extension lets agents play on a two dimensional grid. This makes the problem different from the standard IPD game. When one agent is close to another agent, they play one round of the Prisoner's Dilemma. Each agent can then choose to cooperate to give the other agent a higher chance of procreating, while taking a small penalty to their own chance of procreating. In this extension agents do not have memory, and therefore all interactions between agents can be seen as single rounds of the Prisoner's Dilemma.

Reputation is the second extension. Without reputation each agent does not know anything about the history of its opponent. This means that in a grid where agents only play eachother once a round agents that always defect have free play. When reputation is introduced defection in earlier rounds can result in a negative reputation of that agent, and thereby lower its chance to survive over time.

The aim of this paper is to find out how reputation influences the resistance to bad agents in the Iterated Prisoner's Dilemma. This is done by simulating different configurations without reputation, and comparing them to the same configuration with reputation.

In the next paragraph the related work will be discussed. Section 2 will contain all formal definitions of the models. Next, in section 3 the setup and results of the experiment will be shown. Section 4 contains the reproducability of this research. Section 5 describes the differences between the found results and the known results from earlier papers. Finally section 6 contains the conclusion and recommendations for further work.

1.1 Related work

This section will dive into all the research that has been done on reputation in the Iterated Prisoner's Dilemma.

Yao & Darwen [8] researched reputation in the Prisoner's Dilemma. In that paper reputation of an agent holds the value

1 or -1. It was found that reputation can mitigate mutual defection in IPD's with shorter game length.

Nowak & Sigmund [5] also researched reputation. In their paper the reputation starts at the value 0, and this reputation will go up or down depending on the decision of the agent.

The paper by Chong & Yao [3] informs the reader how the amount of choices in the IPD can change the outcome. In that paper reputation is used to estimate behaviours of future opponent agents. The reputation is a parameter for each strategy.

At last B.Baranski et al [2] have researched reputation in a very similar manner. It is researched how different configurations of starting parameters for reputation influence the outcome. Here it was found that reputation encourages cooperation, and that group reputation encourages cooperation between groups, but discourages cooperation inside a group.

In this research reputation is set to be the percentage of cooperation. This makes it different from how Yao & Darwen [8] and Nowak & Sigmund [5] use reputation. Chong & Yao [3] have done research using reputation, and last B.Baranski et al [2] researches the setup for reputation itself. In this paper however a setup of reputation is tested for its resilience to 'bad agents'.

2 Definitions of relevant terms

In this section all terms will be defined. At first the environment will be explained, afterwards the definition of bad agents will be given, and at last the notion of reputation will be stated.

2.1 Environment

The iterated prisoner's dilemma can be modeled in a spatial configuration. Agents are spawned on an edgeless grid (the top connects to the bottom and the right connects to the left) with a horizontal and vertical side of size G_{size}. The neighborhood of an agent is defined as all cells that can be reached by that agent with one horizontal or vertical move. Each iteration has 4 phases: immigration, interaction, reproduction and death. At first during the immigration phase a random agent is placed on a free cell. During the interaction phase each agent plays one round of the Prisoner's Dilemma with a horizontal or vertical neighbor. If an agent decides to cooperate, they do that at a cost γ, and the other agent will gain δ. In the reproduction phase each agent then has a probability to reproduce PTR. The mutable parameters of the child have a probability p_{mut} to mutate. At last each agent has a probability to die p_{die} in the death phase. This cycle can be seen in figure 1

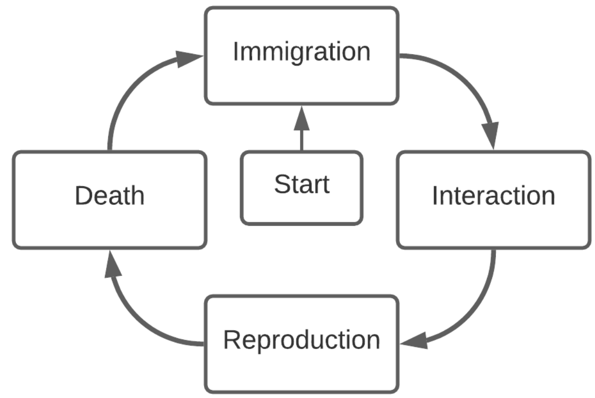

Figure 1: Schematic diagram of the cycle

The two different sets of parameters used can be found in table 2. and are the same parameters used by Axelrod and Hammond[1]. In general (without reputation and bad agents) each agent is set to always cooperate.

	G_{size}	γ	δ	PTR	p_{mut}	p_{die}
easy	50	0.01	0.03	0.12	0.005	0.10
harsh	50	0.02	0.03	0.12	0.005	0.10

Table 2: Parameter settings, from Axelrod and Hammond(2003)

2.2 Bad agents

In general a bad agent is defined as an agent that always defects with other groups, and always cooperates within the same group. In an experiment without groups a bad agent will always defect, since all other agents are not of the same group. In an experiment with groups however, a bad agent will only defect with all agents not in its group.

To implement this, each time a new agent is created (either by immigration or by procreation) that agent has a probability p_{bad} to become a bad agent. This does not influence the offspring of that agent however.

To measure the resilience to bad agents, three different configurations are tested (table 3). The first configuration k_1 is the control configuration. In this configuration no bad agents will be created. The second and third configuration (k_1 and k_2) both introduce bad agents into the population.

Configuration	p_{bad}
k_0	0
k_1	0.05
k_2	0.10

Table 3: Different configurations for p_{bad}, the probability to create a bad agent

2.3 Reputation

In the next two paragraphs both kinds of reputation will be defined. These paragraphs will explain reputation for the standard case and reputation in the case of groups respectively.

Solo-agent-reputation

Each agent has a reputation ρ and a reputation threshold τ. The reputation of each agent is set to be the percentage of cooperation so far. The reputation of an agent a can be written as the amount of games the agent has cooperated in c divided by the total amount of games played t.

$$\rho_a = \frac{c(a)}{t(a)}$$

When ρ_a is close to zero, agent a has defected almost every round. Conversely, when ρ_a is close to one, agent a has cooperated almost every round.

τ is the threshold for cooperation. If one agent a with $\rho = 0.4$ encounters another agent b with $\tau = 0.5$, agent b will choose to cooperate, because $\tau_b \geq \rho_a$. τ is only altered with mutation. Since mutation only occurs when a child is created, the τ value for an agent will be the same its whole life.

Group-reputation

In the case of group-reputation, groups are defined. Each agent belongs to a group. Like the solo-agent-reputation, each agent has a reputation which shows its cooperation percentage. However instead of using this reputation in the IPD game, the group-reputation is used. This is the average of all reputations of the agents in the group, and is calculated as follows where g is a group and n is the number of agents in a group:

$$\rho_g = \frac{1}{n_g} \sum_{a=1}^{n} \rho_a$$

When playing the IPD game each agent has a threshold towards each group. When two agents play one round of the prisoner's dilemma game, the threshold of the agent towards the opponent's group will be checked against the reputation of the opponent's group. Again, if the reputation is higher than or equal to the threshold, the agent will cooperate.

Groups themselves do not influence the decision of the agents directly. It only influences what reputation the agent has (since it is not only its own reputation anymore, but the average of the whole group), and what threshold the opponent uses. This means that when an agent meets another agent of the same group, it could still defect.

3 Experimental Setup and Results

In the next paragraphs the configurations for the experiments will be discussed. After that the hypothesis will be posed. At last the results of the experiments will be examined.

3.1 Experimental configurations

Since there are two configurations for the cooperation-cost γ, two configurations for using reputation, two configurations for using groups and three configurations for the amount of bad agents in the population, there will be 24 different tests in total, as can be seen in table 4. We can then compare the results from the experiments without reputation with the results from the experiments with reputation.

Without reputation			With reputation		
Solo	Low γ	k_0	Solo	Low γ	k_0
		k_1			k_1
		k_2			k_2
	High γ	k_0		High γ	k_0
		k_1			k_1
		k_2			k_2
Group	Low γ	k_0	Group	Low γ	k_0
		k_1			k_1
		k_2			k_2
	High γ	k_0		High γ	k_0
		k_1			k_1
		k_2			k_2

Table 4: All different experiment configurations

The experiments are split into four categories. These consist of the combinations of group-usage and low/high γ. These categories and their names are shown in table 5.

Solo	Low γ	m_1
	High γ	m_2
Group	Low γ	m_3
	High γ	m_4

Table 5: Categories of experiments with their names

Each experiment lasts for 2000 iterations. During the experiment four different variables are tracked if no groups are used, and five different variables are tracked if groups are used. The total amount of agents *pop*, bad agents *bad* and good agents *good* after 2000 iterations will be tracked. The total cooperation percentage is also important. When groups are used the cooperation percentage is split in a cooperation percentage within groups $coop_s$ and a cooperation percentage between agents of different groups $coop_d$. When no groups are used only a total cooperation percentage *coop* is needed.

All tests are ran twelve different times, and the results are shown as the mean over these 12 runs.

3.2 Hypothesis

In configuration k_0 nothing should change. No bad agents are introduced, so agents in tests both with and without reputation should almost always cooperate. Therefore it is expected that all measured variables stay the same or change slightly.

Configurations k_1 and k_2 introduce bad agents. Since reputation punishes defecting, it is expected to see less defection when reputation is included in the test. This in turn should lead to less bad agents in the population. When there are less bad agents, this opens up spots for more good agents. So in total it is expected *pop* to stay the same, *bad* to decrease and both *good* and *coop* to increase when reputation is introduced.

3.3 Results

Experiments without groups

The results of model m_1 can be seen in table 6. As expected, for k_0 it does not matter if reputation is in play or not. As for

k_1 and k_2, the population, amount of good agents and cooperation rates do not differ very much, however the amount of bad agents is reduced by a significant amount.

	pop	*bad*	*good*	*coop*
k_0	1765	0	1765	100.0%
k_0 rep	1760	0	1760	100.0%
δ	-0.3%	0.0%	-0.3%	0.0%
k_1	1753	22	1732	98.9%
k_1 rep	1757	8	1749	99.5%
δ	0.2%	-63.6%	1.0%	0.6%
k_2	1775	26	1749	98.5%
k_2 rep	1759	12	1747	99.2%
δ	-0.9%	-56.2%	-0.1%	0.8%

Table 6: The results of the experiments with no groups and low cooperation cost m_1

The results of model m_2 can be seen in table 7. Just like in m_1, the difference in k_0 is small. The difference in population changes a bit between k_1 and k_2, however in both configurations the amount of bad agents decreases significantly and the amount of good agents increases significantly. The cooperation percentage also increases.

	pop	*bad*	*good*	*coop*
k_0	1444	0	1444	100.0%
k_0 rep	1407	0	1407	100.0%
δ	-2.6%	0.0%	-2.6%	0.0%
k_1	1496	151	1237	89.6%
k_1 rep	1379	66	1312	94.9%
δ	-7.8%	-55.9%	6.1%	5.9%
k_2	1304	263	1041	80.3%
k_2 rep	1421	63	1358	95.5%
δ	9.0%	-76.2%	30.5%	18.8%

Table 7: The results of the experiments with no groups and high cooperation cost m_2

Experiments with groups

In table 8 the results for model m_3 are shown. Again the values for k_0 are very similar. The population difference in k_1 and k_2 is minor, but again the amount of bad agents is reduced, and the amount of good agents is increased. The cooperation rate inside a group decreases slightly, but the cooperation rate between groups increases by a lot.

	pop	*bad*	*good*	$coop_s$	$coop_d$
k_0	1780	0	1780	100.0%	100.0%
k_0 rep	1781	0	1781	99.6%	97.9%
δ	0.1%	0.0%	0.1%	-0.4%	-2.1%
k_1	1707	1346	361	100.0%	28.0%
k_1 rep	1688	846	842	97.3%	43.4%
δ	-1.1%	-37.1%	133.3%	-2.7%	55.1%
k_2	1674	1445	229	100.0%	19.3%
k_2 rep	1686	1241	445	97.9%	25.1%
δ	0.7%	-14.2%	94.6%	-2.1%	30.2%

Table 8: The results of the experiments with groups and low cooperation cost m_3

At last the results for m_4 in table 9. It is notable that the values with and without reputation for k_0 differ. This is not as expected. For k_1 and k_2 the population drops a small amount, however overall there are less bad agents and a lot more good agents. Again the cooperation inside a group decreases slightly, and the cooperation between groups increases.

	pop	*bad*	*good*	$coop_s$	$coop_d$
k_0	1440	0	1440	98.8%	96.5%
k_0 rep	1399	0	1399	98.0%	90.7%
δ	-2.8%	0.0%	-2.8%	-0.9%	-6.1%
k_1	1340	1185	155	100.0%	15.2%
k_1 rep	1283	865	418	93.4%	28.2%
δ	-4.2%	-27.0%	169.5%	-6.6%	85.3%
k_2	1387	1243	144	100.0%	15.8%
k_2 rep	1286	1241	306	92.5%	20.8%
δ	-7.3%	-0.2%	112.7%	-7.5%	31.4%

Table 9: The results of the experiments with groups and high cooperation cost m_4

Conclusion of results

mean δ *pop*	mean δ *bad*	mean δ *good*	mean δ *coop*
-1.4%	-27.5%	45.2%	1.3%

Table 10: Mean differences of the parameters over all experiments

Overall the hypothesis is partially correct. The population does indeed stay the same, albeit with lots of fluctuations (mostly for high γ) The amount of bad people indeed decreases, the amount of good people increases, and the cooperation for m_1 and m_2 increases. It is interesting to note that in model m_3 and m_4 the cooperation within the group decreases due to reputation, but the cooperation between groups increases.

4 Reproducability

The experiments in these paper can be easily reproduced. The experiments are modeled in Netlogo [6], and created using the Ethnocentrism model made by U. Wilensky [7]. The source code for the experiments as well as all the results are uploaded online. [4]

5 Discussion

In this section the results will be compared to the results of B.Baranski et al [2]. Afterwards the definition of bad agents will be discussed.

In the paper by B.Baranski et al [2] a few different conclusions are drawn. Only the conclusions related to the experiments are discussed here. At first it was found that solo-agent-reputation encourages cooperation. In table 6 and table 7 it is shown that the cooperation does not increase significantly in the easy environment, but does increase significantly in the harsh environment. This leads to the second conclusion: "Impact of reputation on cooperation is more obvious when costs increase". This conclusion is only aimed at the solo-agent-reputation, however this also applies to the results of the group-reputation experiments, as can be seen in table 8 and table 9. At last it was concluded that "Group-reputation encourages cooperation with outgroup while discouraging cooperation with in-group." This conclusion fits the experiments exactly. This is also briefly mentioned at the end of section 4.

A bad agent in the group-reputation environment is defined as an agent that only cooperates with its own group. Research was also done into having 'very bad agents' (agents that never cooperate). It was however found that these very bad agents die out very quickly in all experimental configurations, and are therefore not useful for researching the influence of reputation against bad agents.

6 Conclusions and future work

In this paper the resistance to bad agents in the IPD was investigated. At first two forms of reputation were introduced. This was then tested for both an easy (low γ) as well as a harsh (high γ) environment. From the results some clear conclusions can be drawn.

1. Adding reputation has a positive influence on the resulting amount of good people and the total amount of cooperation in all experiments.
2. Adding reputation also reduces the amount of bad agents in the population.
3. The total population does not always increase when using reputation

In total this means that reputation does indeed improve the resistance against bad agents.

Further studies can examine the reason why the cooperation with the own group decreases but the cooperation with other groups increases when reputation is introduced.

Also the reason for higher fluctuations when γ is higher is not known yet.

At last, this paper only considers the basic strategies of 'always cooperate', 'always defect' and 'only cooperate with own group'. Further research could be done on this subject where other strategies are considered.

References

[1] Robert Axelrod, Ross Hammond, Lars-Erik Cederman, Johan van der Dennen, David Lazer, Richard Lenski, Robert Putnam, The Bach, Carl Simon, Michael Cohen, John Holland, Scott Page, Mercedes Pascual, and Rick Riolo. The evolution of ethnocentric behavior. 05 2003.

[2] B. Baranski, T. Bartz-Beielstein, R. Ehlers, T. Kajendran, B. Kosslers, J. Mehnen, T. Polaszek, R. Reimholz, J. Schmidt, K. Schmitt, D. Seis, R. Slodzinski, S. Steeg, N. Wiemann, and M. Zimmermann. The impact of group reputation in multiagent environments. In *2006 IEEE International Conference on Evolutionary Computation*, pages 1224–1231, 2006.

[3] Siang Yew Chong and Xin Yao. Multiple choices and reputation in multiagent interactions. *IEEE Transactions on Evolutionary Computation*, 11(6):689–711, 2007.

[4] Per Knops. IPD resistance using reputation. https://doi.org/10.5281/zenodo.5034734, June 2021.

[5] Martin Nowak and Karl Sigmund. Evolution of indirect reciprocity by image scoring. *Nature*, 393:573–7, 07 1998.

[6] Uri Wilensky. Netlogo. http://ccl.northwestern.edu/netlogo/, Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL, 1999.

[7] Uri Wilensky. Netlogo models library: Ethnocentrism. https://ccl.northwestern.edu/netlogo/models/ethnocentrism, Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL, 2003.

[8] Xin Yao and Paul Darwen. How important is your reputation in a multi-agent environment. 07 2001.