Cooperating with Trusted Parties Would Make Life Easier

Pasquale Caianiello, Stefania Costantini, Giovanni De Gasperis(✉), and Subhasis Thakur

Dipartimento di Ingegneria e Scienze Dell'Informazione e Matematica, Universitá Degli Studi Dell'Aquila, Via Vetoio 1, 67100 L'Aquila, Italy
{pasquale.caianiello,stefania.costantini,giovanni.degasperis, subhasis.thakur}@univaq.it

Abstract. We experimentally analyze the performance of a heterogeneous population of agents playing the Iterated Prisoner's Dilemma with a possible prior commitment ad a posterior punishment for defection. We argue that the presence of agents with a probabilistic strategy that depends on trust and reputation enforces a better performance of typically cooperative agents.

1 Introduction

This paper deals with cooperation enforcement in the Prisoner's Dilemma (PD) game. Though the PD is a non-cooperative game, in several application fields, ranging from biology to social sciences to multi-agent systems (MAS), it has become the leading paradigm to model and discuss cooperative behavior. In one-shot Prisoner's Dilemma, two players simultaneously decide to either cooperate (C) or defect (D). If both play C, they get more than if both play D, otherwise in case one defects and the other cooperates, then the defector gets the highest payoff, while the cooperator gets the lowest. Consequently, rational choice would imply that it is safer for each player to defect, even though both would get a better payoff in case of cooperation. The situation were both players do not cooperate is the Nash equilibrium of the PD game. Several approaches for promoting cooperation have been proposed, some introducing for instance voluntary rather than compulsory participation [1] with punishment for non cooperating agents, some others introducing prior commitments and possibly posterior punishment for non cooperation [2]. The main aspect to consider is that both a-priori negotiation for reaching a commitment and a-posteriori administration of punishment have costs.

As a matter of fact, negotiation and punishment are complementary in the way they try to induce cooperation: prior commitments function better with "compliant" agents, and punishments with "free riders" which pursue their momentarily best interest. Commitment definition and formation has been extensively studied (cf. [3] and the references therein) and both mechanisms have been considered in the field of software agents and multi agent systems

M. Gavanelli et al. (Eds.): AI*IA 2015, LNAI 9336, pp. 128–135, 2015.
DOI: 10.1007/978-3-319-24309-2_10

(MAS) where commitment are usefully employed in many fields that include inter-agent communication. Prior commitment, though costly, may be applied on a probabilistic basis [4]. Recent studies (see, e.g. [5]) discuss the conditions when a strategy that combines the two mechanisms is better than either strategy by itself in a MAS. It has been advocated in more general terms [6] that the tendency to making prior agreements rather than just requiring a posteriori compensations emerges from a variety of examples in biological and social contexts, thus suggesting that this behavior could have been shaped by natural selection, and, therefore, "good agreements make good friends" [6].

In this paper, we cope with the Iterated Prisoner's Dilemma, where players engage in Prisoner's Dilemma repeatedly and change their strategy according to a shared indicator built upon previous actions of all involved agents. We assume the implicit existence of a game manager that provides payments for game payouts and collects punishment fines. In such a research setting, both a-priori commitments and a-posteriori punishment for defeating commitments have been used in the existing literature in support of "apology" (see [7,8]). We propose to use forms of public trust evaluation as indicators for implementing strategies eventually leading to more successful course of actions. In particular, we argue that long term gain would result from effective commitment reached with trusted parties, where trust evaluation evolves dynamically with game repetitions. We provide results of computational simulations showing that the adoption of trust evaluation (cf. [9,10] and the reference therein for a discussion about the notion of trust and of trust-update mechanisms) enforces a higher final gain for agents playing the iterated PD.

1.1 Game Theory Basic Notions

Game theory is a study of strategic decision making involving cooperative and non-cooperative agents. This paper falls into the field of non-cooperative game theory as we study how a group selfish agents make decisions that maximize their respective utility. Nash equilibrium is a way to model the equilibrium of such decision making process. In this paper agents decisions are based on the evaluations of the opponent, by means of trust and reputation. While in cooperative game theory we study how a group of agents can decide on the rules of cooperation using their respective share of the utility gained from such cooperation, we argue that the evaluation of trust and reputation may be a viable way to promote cooperation as concepts from both cooperative and non-cooperative game theory may be used together in the study of multi-agent decision making.

2 Model and Methods

As we are addressing the question of commitment in the Prisoner Dilemma, we consider the scenario where the game is played in several iterated rounds among a non homogeneous population of different agents. In each round two players are selected at random with uniform probability to play a one-shot game.

Before making their choice (C or D), either player may simultaneously propose a commitment to cooperation which the other player may accept or deny. If both players propose, then commitment is established; if only one proposes, then commitment is established only if the other accepts; finally if either does not accept or neither proposes then no commitment is established. Commitment proposal has a cost ϵ that is shared in equal parts if the commitment is established. On the other end, if no commitment is established, the entire cost is charged only to the proposing player, if there is one.

After the commitment proposal/accepting preliminary move, the players make their simultaneous choice C or D, they get their resulting payoff and, if either agent plays D in a round when a commitment was established, it will have to pay a penalty δ to the deceived opponent. In Fig. 1 we show the payoff matrix where $T > R > P > S$.

	C	D
C	R,R	S,T
D	T,S	P,P

Fig. 1. Payoff matrix

We present simulation results by computing agents cumulative wealth obtained as a net outcome from rounds payoff, commitment, and penalty payments.

2.1 Players Typologies and Profiling

The two-moves version of the prisoner dilemma game allows the definition of typical agents with deterministic behaviors. We consider a population of several different playing agents of two major classes as described in Fig. 2. Agents in the first class behave according to a strategy that does not change over time, the names associated to their behavior are already established in related literature, [5]. We describe them with minor differences and some new entries, in particular the BASTARD agent who always tries to establish a commitment but afterwards always deceives, and the SCHIZO agent, who always tries to establish a commitment but afterwards behaves inconsistently by playing D when the commitment is established and plays C when there is no commitment.

Among the all theoretically possible agent behaviors, the only missing is the one that never proposes or accepts and then plays C anyway. In fact although it seems that there are 2^4 possible deterministic behaviors , it should be noted if an agent proposes, then according to the game rules it will always accept. So no player can always propose and never accept.

In the bottom lines of Fig. 2 we describe a class of agents whose strategy is probabilistic an depends on the global profiling of their opponent represented as trustworthiness θ measuring the agent's disposition to comply to commitment, and reputation ρ, measuring the agent's willingness to play C role. Such profiling is globally updated at each round with the simple reinforcement rule $x(t) := x(t-1) + \Delta x$ that increases (or decreases) their trustworthiness and reputation, x is θ or ρ, by a fraction Δ of what they miss to get to the maximum (or minimum):

$$\Delta\theta = \begin{cases} +\alpha(1-\theta) \text{ } \mathit{if\ commit\ and\ play\ C} \\ -\alpha\theta \text{ } \mathit{if\ commit\ and\ play\ D} \\ 0 \text{ } \mathit{if\ no\ commit} \end{cases}$$

$$\Delta\rho = \begin{cases} +\alpha(1-\rho) \text{ } \mathit{if\ play\ C} \\ -\alpha\rho \text{ } \mathit{if\ play\ D} \end{cases}$$

where $0 < \alpha < 1$ and drives the rate of change of θ and ρ during subsequent rounds.

The class of agents with probabilistic behavior that we consider include: RANDOM who in any game and with an opponent just flips a coin to decide what to do, and the others, but DIPLOMAT, play C with probability that is equal to the opponent reputation ρ if no commitment is established, and modulate their moves in establishing commitment with a probabilistic choice depending on the opponent profiling. In particular TRUSCoop who decides whether to establish a commitment with a probability θ, always plays C if a commitment is established; TRUST who proposes a commitment with probability θ, never accepts commitments and plays C with probability θ if a commitment is established;

name	propose	accept	coop on commit	coop on no commit
C	always	always	always	always
D	never	never	*irrelevant*	never
COMP	always	always	always	never
FAKE	never	always	never	never
FREE	never	always	always	never
BASTARD	always	always	never	never
SCHIZO	always	always	never	always
RANDOM	P=1/2	P=1/2	P=1/2	P=1/2
TRUST CooP	P=θ	P=θ	always	P=ρ
TRUST	P=θ	always	P=θ	P=ρ
REP Only	never	never	P=ρ	P=ρ
DIPLOMAT	always	always	P=θ∗ρ	P=ρ

Fig. 2. The probability P of playing C for agents in the simulation, yellow are pure deterministic agents, orange are probabilistic.

REPonly who never commits, and DIPLOMAT who always tries to establish a commitment and, when it is established than plays D with a probability that is equal to the product of θ and ρ, and otherwise plays D when the commitment is not established.

We stress the fact that in the present work we deliberately kept "adaptive" agents strategies simply depend on the global profiling of agents trustworthiness and reputation, in order to gain a preliminary insight on possible outcomes where trust and reputation are involved in agents decisions. Obviously more complex adaptive strategies taking into account different aspects of agents behavior may be conceived of, and will be the object of sequel work.

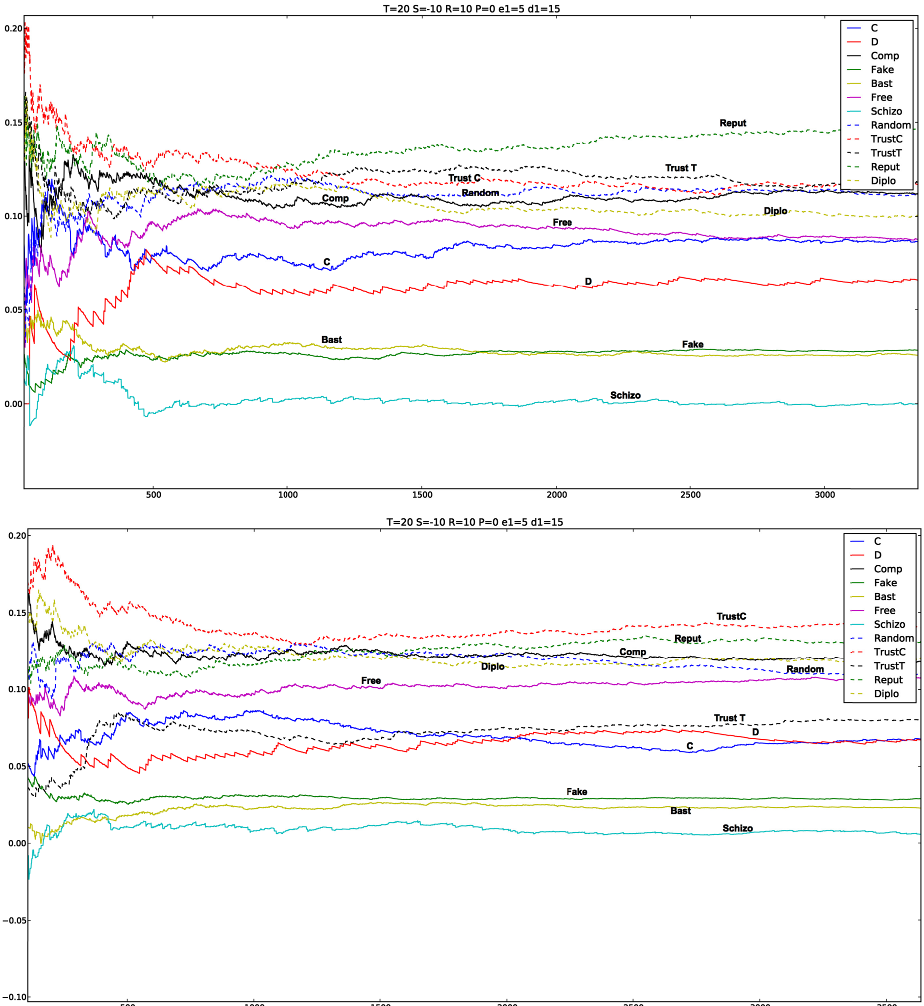

Fig. 3. Trend of agents relative wealth in two different simulations with same parameters values.

3 Simulation Results

We have performed a number of experiments that we show and discuss below. The outcome is that the introduction of trust and reputation increases the level of cooperation while decreasing the cost for both single agents and overall MAS. Each simulation is initialized with a random population of 1000 agents chosen with uniform probability among the 12 agent described in Fig 2. Each simulation is run for 10000 rounds where two players are chosen at random uniformly in the population.

3.1 Agents Performance

In Fig. 3 we plot the relative agents wealth in two different simulations, with the same choice of ϵ and δ and other parameters. We noticed that different runs of the simulation with same parameters do lead to somewhat different results in agents relative performance. That, in fact, it's due to the complexity of the game that we are simulating and testifies for a dependency of the final result on the initial random choices. Chance and luck do play a role in the Iterated Prisoner Dilemma.

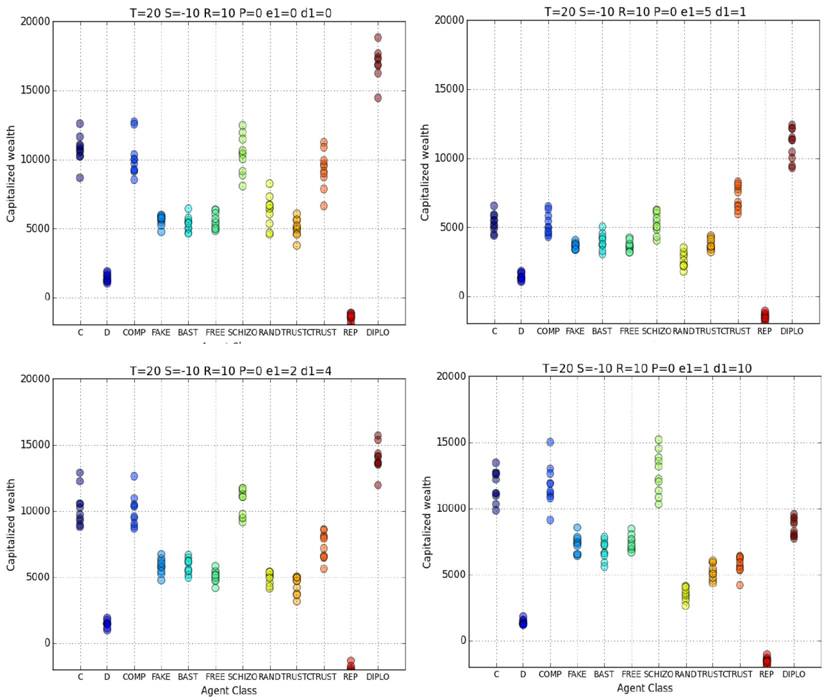

Fig. 4. Simulation results for different choice values of the commitment cost ϵ (e1) and penalty δ (d1).

To overcome this problem in order to arrive at more definite relative evaluation of agents performance we decided to get a better idea by looking at ten different simulations and look at the cluster of final results that the agents obtained. In Fig. 4 we report about the final wealth obtained when running ten simulations of the game with different values of ϵ and δ, as obtained in the simulations and presented in order of increasing values of the ratio δ/ϵ. As we see there is a definite best performance of the DIPLOMAT, except for high values of δ when the DIPLOMAT performance suffers for its probabilistic choice, and consistently cooperative agents as C and COMP perform better. Notice that SCHIZO does have an appeal in almost any situations.

3.2 Agents Wealth vs Commitment Cost and Punishment

In order to appreciate the possible influence of the chosen values of ϵ and δ, in Fig. 5 we plot the average wealth obtained by all the agents of the same type against a few chosen values of ϵ and δ plotted for increasing δ/ϵ

We argue that a combination of the commitment cost plus violator's punishment with a trust mechanism, where involved agents are keener to pay commitment cost if dealing with trusted agents, should result in better agents performance. Moreover, the fact that the population includes different types of agents, many of which behave according to a probabilistic strategy depending on trust and reputation, also modifies relative performance of agents with a fixed behavior.

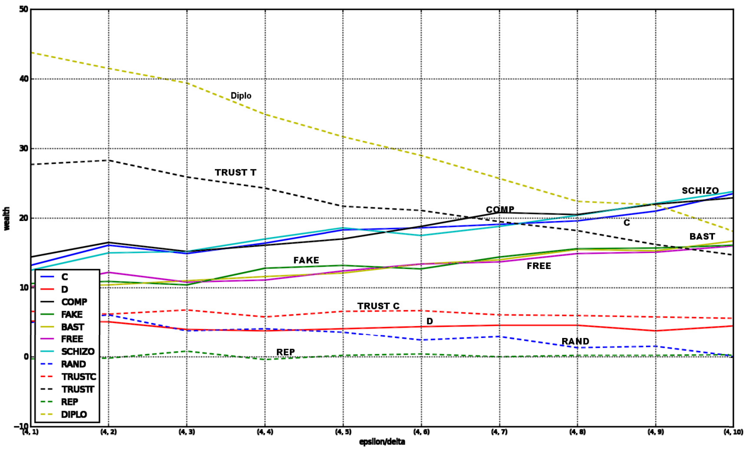

Fig. 5. Agents average wealth for choice values of ϵ (e1) and δ (d1) plotted for increasing (ϵ, δ)

4 Concluding Remarks

In this paper, we have advocated trust evaluation to promote cooperation in the Iterated Prisoner's Dilemma with prior commitment, and performed simulations with a mixed population of deterministic and probabilistic agents whose move depends on a simple profiling of opponents trustworthiness and reputation. The experimental results suggest that some probabilistic agents relying on trustworthiness and reputation perform consistently better, though their performance decreases for high values of the ratio of punishment over commitment costs in favor of more consistent typically cooperative agents. We can argue that, on the one end trust and reputation provide a solid playground for agents to achieve a better payoff in cooperation games, on the other a reasonable penalty for defecting commitments could promote cooperation.

References

1. Hauert, C., Szabó, G.: Prisoner's dilemma and public goods games in different geometries: Compulsory versus voluntary interactions. Complexity **8**(4), 31–38 (2003)
2. Pereira, L.M., Santos, F.C., et al.: The emergence of commitments and cooperation. In: Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems, vol. 1, pp. 559–566. International Foundation for Autonomous Agents and Multiagent Systems (2012)
3. Singh, M.P.: Commitments in multiagent systems: some history, some confusions, some controversies, some prospects. In: Paglieri, F., Tummolini, L., Falcone, R., Miceli, M. (eds.) The Goals of Cognition: Essays in Honor of Cristiano Castelfranchi, pp. 601–626. College Publications, London (2012)
4. Chen, X., Szolnoki, A., Perc, M.: Probabilistic sharing solves the problem of costly punishment. New Journal of Physics **16**(8) (2014)
5. Han, T.A., Lenaerts, T.: The efficient interaction of costly punishment and commitment. In: Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2015, Istanbul, Turkey, May 4–8, 2015, pp. 1657–1658 (2015)
6. Han, T.A., Pereira, L.M., Santos, F.C., LenaertsJ, T.: Good agreements make good friends. Scientififc Reports **3** (2013). (Online, Open-Access)
7. Han, T.A., Pereira, L.M., Santos, F.C., LenaertsJ, T.: Why is it so hard to say sorry? evolution of apology with commitments in the iterated prisoner's dilemma. In: Rossi, F. (ed.) IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3–9, 2013, IJCAI/AAAI (2013)
8. Okamoto, K., Matsumura, S.: The evolution of punishment and apology: an iterated prisoner's dilemma model. Evolutionary Ecology **14**(8), 703–720 (2000)
9. Castelfranchi, C., Falcone, R.: Trust theory: A socio-cognitive and computational model, vol. 18. John Wiley & Sons (2010)
10. Wang, Y., Singh, M.P.: Formal trust model for multiagent systems. In: Veloso, M.M. (ed.) IJCAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence, pp. 1551–1556 (2007)