Research Plan for New Generation Iterative Prisoner's Dilemma

Martin Toman

23 April 2021

Title: Keep your enemies closer and be loud about it

Author: Martin Toman

Responsible Faculty: Dr. Neil Yorke-Smith

Peer group members: Roberta Gismondi, Per Knops, Raymond Timmermans, Tommaso Tofacchi

Abstract

Under what conditions is cooperation the optimal strategy? To answer this question a two-player exchange game was defined: The Prisoner's Dilemma. However, to sustain cooperation over a long period an external mechanism is required. It was shown that keeping track of the reputation of all participants (e.g. Ebay's seller rating) can greatly boost levels of cooperation in the population. We propose a new research into the effects of reputation in a Spatial version of the game using only local interactions to observe and exchange the reputation information. This design more closely fits decentralized systems and a similar model has already been used to improve peer-to-peer protocols.

Background of the research

How can we encourage and sustain cooperation? Humans dominate their environments thanks to our ability to cooperate flexibly and at scale, as argued by Harari [8]. To study the conditions necessary for cooperation to flourish we will need a suitable model of an activity with temptations to defect and punishment for doing so.

In 1950, Albert Tucker named a particular two-player exchange game "The Prisoner's Dilemma" [10]. This game elegantly captures the difficulty of the decision between cooperation and defection in a single choice. Despite being so simple compared to the complexity of the problem it is representing, it was used to model many aspects of behaviour in systems of selfish individuals; and, as formulate by Axelrod [1], for "discovery of the precise conditions that are necessary and sufficient for cooperation to emerge".

In the case of a one-off exchange, there being no opportunity for a follow-up punishment, the rational behaviour is defection. (This extends to all rounds for a fixed-length game, inductively [1].) The truly interesting behaviour arises if there is no end; or, at least, if there is no way for the participants of the game to know when or even if there is an end. One has to expect that even a single defection can be infinitely punished by never again cooperating with the culprit [6]. Such a risk may just not be worth it.

The defectors can, naturally, only be punished if they can be identified and known to others. This is why services like Ebay or Airbnb have a rating system in place. Presence of a reputation system has been shown to strongly boost cooperation, as shown by Stahl [12] and Camera and Casari [2]. These studies used groups of volunteers as game participants and explored the effects of various information being public - varying from only the latest move of the current opponent, to full histories of all moves taken by every participant.

Some limitations of these studies were the fact that they used humans as game participants and were thus limited to relatively small groups with few rounds; they also used external infrastructure for information passing: therefore eliminating noise, delays, and deliberately wrong information. As shown by Gevers and Yorke-Smith [7], not all strategies that perform well in noise-less environments can do so under the presence of noise.

Using external infrastructure for passing information also meant that the transmission speed was uniform for all participants receiving all necessary information in time for their next round of the game. These are non-negligible idealizations; relaxing them would yield a model closer to real-world systems and could change the results drastically. There is no way to tell without testing it out.

Research Question

Can we sustain cooperation with only locally determined reputation? Presence of a global reputation "improves the frequency of cooperators in the population" [5], but has problems of its own and as a centralized system is vulnerable to censorship and malfunctions can be very costly. By decentralizing the reputation system, we can leverage the participants themselves to create conditions ideal for cooperation while avoiding many shortcoming of the centralized system. This is interesting in the context of peer-to-peer systems, where a centralized system is not a tenable solution as it would counteract most of the benefits of a peer-to-peer solution. Jun and Ahamad [9] have successfully used an Iterated Prisoner's Dilemma model to investigate the incentive mechanism in BitTorrent [4], a peer-to-peer file distribution protocol.

To answer the research question, we present the following sub-questions to serve as an outline for leading us to the final answer; all while exploring interesting interactions along the way.

1. Can we replicate the design and results of Smaldino et al. [11]?
2. How can we extend the model with variable-length memory and what strategies can we define on this new model? What are the effects of shorter/longer memory and what is the optimal length?
3. What is the best way to combine own observations and gossip from peers? We want to combine the data in such a way that we keep information about participants we are most likely to interact with in the next round.
4. Does openly gossiping with other participants boost cooperation, when compared to a no-gossip scenario?
5. At what range is the gossiping most effective? Is it enough to exchange information with other participants on contact (i.e. on the same distance that a round can be played)?
6. What is the difference between gossip-then-play, and play-then-gossip models? Does it make a difference if we gossip before or after playing the round?
7. Can we maintain this cooperation level by only gossiping compressed categorical data? (e.g. "hostile"/"neutral"/"friendly" instead of "defected in 82% of rounds when unprovoked") So as to model the way natural language works and how local context skews the absolute meaning of words: "hostile" for one participant could be "friendly" for another.

Method

To explore the effects of local reputation, built via openly gossiping with local peers, we are gonna use a computer simulation of a multi-agent spatial environment; we will base it on the design of Spatial Iterated Prisoner's Dilemma, as used by Smaldino et al. [11]. A single round of the game is defined using a payoff matrix as shown in Table 1, with $T > R > P > S$ and $2R > T + S$ [3].

		Opponent's move	
		Cooperate	Defect
Player's move	Cooperate	Player: R Opponent: R	Player: S Opponent: T
	Defect	Player: T Opponent: S	Player: P Opponent: P

Table 1: Payoff matrix

Participants are agents living in a 2D grid, accumulating energy via repeatably engaging in rounds of PD as defined above. Agents are moving at the same speed, and their order in which they take their turn is randomly determined at every clock tick. They pay a fixed cost to survive to a next round, which is subtracted from all peers at the end of each turn (agents who deplete their energy die and are removed from the simulation). Another use of the energy is to constrain reproduction and allow only the most successful agents to reproduce, producing an offspring with an identical strategy.

We will expand the model by giving the agents a (limited) memory to keep track of past actions of other participants and later allow them to actively and freely share this knowledge by gossiping at various range.

To determine the effectives of gossiping in inducing cooperation, we will observe the rate of convergence to a population of cooperators, stopping the game once stable equilibrium is achieved. We will use the no-gossip model as a benchmark and compare various range at which gossiping is possible to determine the optimal amount of information transmitted necessary for cooperation to win in the population. After that we will compress the gossip into only labels ("hostile"/"neutral"/"friendly") determined by agents based on their local observations.

To achieve the goal of this paper, we define the following subtasks which have to completed:

1. Reimplement the design of Smaldino et al. [11]. (10 hours)
2. Run the model and try to replicate the results (5 hours)
3. Expand the model with memory (10 hours)
4. Determine strategies to be used with the memory model (15 hours)
5. Run experiments with varying memory length (determine optimal memory length) (10 hours)
6. Expand the model with gossip at range (20 hours)
7. Run experiments with varying gossip range (determine optimal range) (10 hours)
8. Locally compress transmitted information based on local observations (10 hours)
9. Run experiments to determine the effects of the compression (15 hours)

(Note: the time estimated are conservative but optimistic approximations.)

Each of these subtasks is dependant on the successful and total completion of all the previous ones. All of them have to completed sequentially in the order as defined above.

For the implementation of the spatial agent model, we are planning to use MESA (`https://mesa.readthedocs.io/en/master/`), a Python framework for agent-based simulations. This has been chosen over other options, because of the popularity of the Python programming language (especially in the context of Jupyter notebooks); we hope this will make this project more approachable to a wider audience and inspire more experimentation or allow others to easier extend/modify the implementation.

Planning of the research project

In this section, a preliminary weekly plan is presented. It is based on the subtasks defined for this research in the previous section and referenced by numbers on the right margin.

Week 0

1. Read preliminary research provided by the responsible professor
2. Explore foundational papers [1, 11]
3. Read last year's student-papers on this topic
4. Research recent papers and new research directions
5. Formulate possible research questions

Week 1 - April 25

1. **(April 19 10:45)**: Kick-off meeting
2. **Deadline (April 19)**: planning week 1
3. **(April 20 17:00)**: Meeting: peer group + responsible professor
4. **Deadline (April 20)**: information literacy 2
5. Setup LaTeX
6. Explore interesting "PD with communication" papers
7. Decide on a research question & narrow down the scope of the paper
8. Find more relevant research: investigate the state of the art
9. **Deadline (April 25)**: research plan

Week 2 - May 2

1. **(April 27 16:00)**: Meeting: peer group + responsible professor
2. Setup Python development environment
3. Replicate Smaldino et al. [11] **(1)**
4. Compare results of the reimplemented model with the original **(2)**
5. Study papers concerning reputation systems in the context of PD [2, 12]
6. Prepare research plan presentation
7. **Deadline (May 2)**: research plan presentation

Week 3 - May 9

1. Add memory to the SPD model
2. Meeting: peer group + responsible professor
3. Add memory to the model **(3)**
4. Define interesting memory-based strategies **(4)**
5. Compare results of the model with memory with the original **(5)**
6. Write background + model design sections of the final paper

Week 4 - May 16

1. **(May 10 10:45)**: responsible research session
2. Meeting: peer group + responsible professor
3. Formulate a communication model to be used for gossip at range
4. How can we combine local with received data so as to keep the most relevant in the memory **(6)**
5. Test a self-learning model instead of a fixed strategy

Week 5 - May 23

1. Compare gossip-then-play with play-then-gossip models **(7)**
2. Meeting: peer group + responsible professor
3. Write method sections of the final paper
4. Document preliminary results in the paper
5. Prepare midterm poster for presentation of progress
6. **Deadline (May 19)**: midterm poster

Week 6 - May 30

1. Meeting: peer group + responsible professor
2. Implement communication compression **(8)**
3. Rerun experiments with communication compression and compare results
4. Test out communication compression effects on the cooperation levels **(9)**
5. Find the optimal levels of compression, which are able to sustain cooperation

Week 7 - June 6

1. Meeting: peer group + responsible professor
2. Format diagrams & plots for the final paper
3. Write results sections & conclusion of the paper
4. Write responsible research section of the paper

Week 8 - June 13

1. Meeting: peer group + responsible professor
2. Write abstract & finalize paper
3. **Deadline (June 7)**: paper draft v1
4. **Deadline (June 10)**: peer review

Week 9 - June 20

1. Meeting: peer group + responsible professor
2. Fix issues with paper & incorporate feedback
3. Create a jupyter notebook for the model & make it available online
4. **Deadline (June 16)**: paper draft v2

Week 10 - June 27

1. Meeting: peer group + responsible professor
2. Incorporate feedback in the paper
3. Outline future work in the paper and add this to the end of the paper
4. Polish the final version of the paper
5. Prepare final poster
6. **Deadline (June 27)**: final paper

Week 11 - July 4

1. Meeting: peer group + responsible professor
2. **Deadline (June 29)**: final poster
3. Poster presentation

References

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

[2] Camera, G. and Casari, M. (2009). Cooperation among strangers under the shadow of the future. *American Economic Review*, 99:979–1005.

[3] Chammah, A., Rapoport, A., Chammah, A., and Orwant, C. (1965). *Prisoner's Dilemma: A Study in Conflict and Cooperation*. Ann Arbor paperbacks. University of Michigan Press.

[4] Cohen, B. (2003). Incentives build robustness in bittorrent.

[5] Dong, Y., Sun, S., Xia, C., and Perc, M. (2019). Second-order reputation promotes cooperation in the spatial prisoner's dilemma game. *IEEE Access*, 7:82532–82540.

[6] Friedman, J. W. (1971). A Non-cooperative Equilibrium for Supergames. *The Review of Economic Studies*, 38(1):1–12.

[7] Gevers, L. and Yorke-Smith, N. (2020). Cooperation in harsh environments: The effects of noise in iterated prisoner's dilemma. *Proceedings of BNAIC/BeneLearn 2020*, pages 414–415.

[8] Harari, Y. (2015). *Sapiens: A Brief History of Humankind*. Harper.

[9] Jun, S. and Ahamad, M. (2005). Incentives in bittorrent induce free riding. In *Proceedings of the 2005 ACM SIGCOMM Workshop on Economics of Peer-to-Peer Systems*, P2PECON '05, page 116–121, New York, NY, USA. Association for Computing Machinery.

[10] Kuhn, S. (2019). Prisoner's Dilemma. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2019 edition.

[11] Smaldino, P. E., Schank, J. C., and McElreath, R. (2013). Increased costs of cooperation help cooperators in the long run. *The American Naturalist*, 181(4):451–463. PMID: 23535611.

[12] Stahl, D. O. (2013). An experimental test of the efficacy of a simple reputation mechanism to solve social dilemmas. *Journal of Economic Behavior & Organization*, 94:116–124.