

Getting recommendation is not always better

Zeynep B. Cinar and Haluk O. Bingol

Abstract—We present an extended version of the Iterated Prisoner’s Dilemma game in which agents with limited memory receive recommendations about unknown opponent to decide whether to play with. Since agents can receive more than one recommendations about the same opponent, they have to evaluate the recommendations according to their disposition such as optimist, pessimist, or realist. They keep their firsthand experience in their memory. Since agents have limited memory, they have to use different forgetting strategies. Our results show that getting recommendations is not always perform better. We observe that realist performs the best and optimist the worse.

Index Terms—Iterated Prisoner’s Dilemma, recommendation, optimist, pessimist, realist, forgetting, multi-agent systems.

I. INTRODUCTION

We live in a complex and crowded world. It is so crowded that it is impossible to “know” everybody [1]. We know only a very small fraction of the population. We operate within this network of known people. As we interact with them, we classify them based on our firsthand experience. Next time we need to interact again, we use this information.

Since bold interaction with an unknown person may not be a good idea, proceeding with firsthand experience is not the only way that we use to make a judgement about somebody. If we do not know a person, we ask for recommendations from people that we already know. We do that all the time in real life.

It is common sense that having recommendation is better. In this work, we investigate whether

having a recommendation is always a better alternative in the context of iterated prisoner’s dilemma game. Our findings indicate that it is not always the case.

II. BACKGROUND

A. Optimism and pessimism

There are different orientations in humans to suggest that what kind of things they could do in a given situation [2]. These orientations have an impact on trust [3]. Marsh presents dispositions in terms of how an agent estimates trust [4]. Optimism, pessimism and realism are the notions of dispositions of trust. Each different disposition results in different trust estimations from an agent. An optimist expects the best in people and is always hopeful about the result of the situations. A pessimist, unlike the optimist, sees the worst in people, always looks at the situations through doubting eyes. While optimism and pessimism are extreme cases, there are other cases such as “realism”.

B. Prisoner’s dilemma

Prisoner’s Dilemma is a two-person game, in which players can choose between two different strategies: “Cooperate” or “Defect”, without knowing their opponent’s choice [5]. As summarized in Table I, if a person cooperates while the other defects, cooperator gets the sucker payoff S and defector gets the temptation payoff T . On the other hand, if both players choose to cooperate, then they both get the reward payoff R . Lastly, in the case of mutual defection, both players get the punishment payoff P . In Prisoner’s Dilemma game, the payoffs should satisfy both $S < P < R < T$ and $S + T < 2R$.

Z. B. Cinar and H. O. Bingol are with the Department of Computer Engineering, Bogazici University, Istanbul, 34342 Turkey e-mail: XXXX@boun.edu.tr.

Manuscript received XX XX 20XX; revised XX XX, 20XX.

TABLE I: Payoff matrix. $(S, P, R, T) = (0, 1, 3, 5)$,

		Player X	
		Cooperate	Defect
Player Y	Cooperate	(R, R)	(S, T)
	Defect	(T, S)	(P, P)

For a single Prisoner’s Dilemma game, it is more advantageous to defect. But, when the game repeats, things change. Iterated prisoner’s dilemma differs from the original concept of a prisoner’s dilemma due to the fact that players can memorize the past interactions of their opponent and can change their strategy accordingly [5], [6]. Players can learn about the behaviors of their opponent and have the opportunity to penalize the agents for previous defective decisions.

Iterated prisoner’s dilemma is used to understand cooperation. Some evolutionary approaches, which promotes cooperation, are investigated in the literature [7], [8], [9]. Some works are based on reputation of the agents, where agents cooperate with the ones that has good reputation [5], [10], [11]. In this paper we propose a non-evolutionary model that also promotes cooperators. We will see that cooperators can get average payoff that is larger than that of defectors.

C. Cetin and Bingol’s model

In the model proposed by Cetin and Bingol, agents play Iterated Prisoner’s Dilemma game in which players can accept or refuse to play with their partner [12], [13]. There are N agents in the population. Agents are not pure cooperators or pure defectors. Agents have an internal parameter of ρ . Agent i cooperates with probability ρ_i , which is called *cooperation probability*. There are two types of agents, one group, called *cooperators*, has a ρ value larger than 0.5. The other group, called *defectors*, has a ρ value, which is less than 0.5.

Decision to Play. Cetin adopts choice-and-refusal rule [14]: If an agent “knows” that the opponent is a defector, then it refuses to play. Otherwise, it plays. That is, two agents, say i and j , are randomly selected and offered to play Iterated Prisoner’s Dilemma game. Both agents evaluate their opponent and decide whether to play or not. (i) If an agent does not know the opponent, then it has to play. (ii) If it “knows” the opponent

as “cooperator”, then it plays. (iii) If i “knows” j as “defector”, then i refuses to play.

Perception. In order agents to “know” each other, agents have some memory so that they can keep track of previous games with the same agent. Suppose agent i plays with agent j . Agent i keeps two numbers in its memory. The number c_{ji} is the number of times that j cooperates and d_{ji} is the number of times that j defects when j plays with i . Then the *perceived cooperation ratio* is defined as

$$t_{ij} = c_{ji} / (c_{ji} + d_{ji}). \quad (1)$$

If the perceived cooperation ratio is larger than 0.5, i considers j as *cooperator*, otherwise as *defector*.

Memory. We assume that each agent has an identical memory capacity of size $M \leq N$, called *memory size*. That is, each agent can keep track of at most M opponents. *Memory ratio*, defined as $\mu = M/N \in [0, 1]$, is the percentage of agents that can be kept in ones memory.

Forgetting strategies. Agent stores each opponent in a different slot in its memory. Eventually, the agent will run out of memory for $M < N$. After that point, to create memory space for a new opponent, the agent has to “forget” a known opponent. There are several forgetting mechanisms investigated in the model. (i) Players prefer to forget cooperators first, denoted by FC. (ii) Players prefer to forget defectors first, denoted by FD. (iii) Players prefer to forget randomly, denoted by FR [12].

III. PROPOSED MODEL

In Cetin and Bingol’s model, if an agent does not know the opponent, it has to play [12], [13]. However, in real life, we use our social network to obtain information about a person that we do not know. We extend the model so that agents can get recommendations about the opponents that they do not know.

A. Perception graph

Perceptions between agents can be represented as a weighted directed graph as given in Fig. 1. In a *perception graph*, agents are represented by vertices. When i plays with j for the first time, two

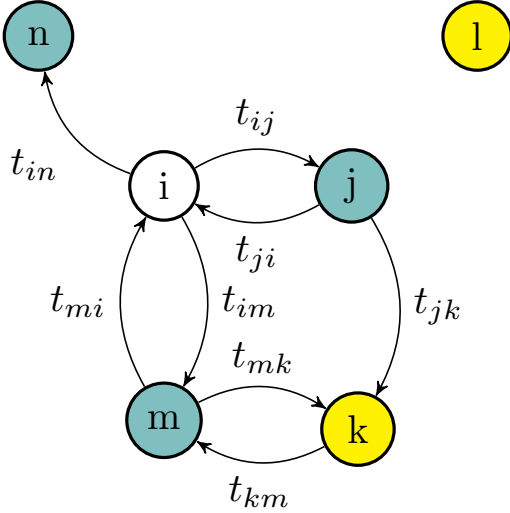


Fig. 1: Perception graph. Agents j , m and n are in the 1-neighborhood of i , while l and k are not.

arcs will be created, namely, one from i to j , and one from j to i . The perceived cooperation ratio of j with respect to i , denoted by t_{ij} , is assigned to the arc from i to j as a weight.

As the game proceeds, symmetric connectivity between agents is broken due to forgetting mechanism, as in the case of j and k . Suppose agents j and k played before. Hence arcs (j, k) and (k, j) were created. Later, when agent k chooses to forget j due to lack of memory space, arc (k, j) and its corresponding weight t_{kj} are removed from the graph and its memory. Note that since j still keeps k in its memory, arc (j, k) and t_{jk} are intact.

B. Decision to play or not

In Fig. 1, suppose i is one of the selected agents to play. As far as the opponent is concerned, there are three cases. (i) Known opponent. For example, j is known by i . Then all i has to do is to check the perceived cooperation ratio. If $t_{ij} > 0.5$, it plays. However, if there is no data about the opponent in its memory, as in the cases of k or l , agent i plays with it in Cetin and Bingol’s model [13]. In our model, agent i asks for “recommendations” from its neighbors. There are two possible cases: (ii) Unknown opponent without any recommendation. If nobody in its neighborhood knows the opponent, i plays with

the opponent as in the case of l . (iii) Unknown opponent with recommendations. Any neighbor, such as j , that knows k , provides its own perception t_{jk} about k . If there were only one such agent j , decision of i would be relatively easy: it will play if $t_{jk} > 0.5$. But usually, there are many such agents, e.g. j and m . Then i has to evaluate conflicting recommendations received from them.

C. Evaluation of recommendations

We define 1-neighborhood of i , denoted by $\Gamma(i)$, as all agents to which there is an arc from i . Note that $\Gamma(i)$ is composed of agents in i ’s memory. Hence if an agent j removed from i ’s memory, it is also removed from $\Gamma(i)$.

The set of recommender agents of i about k is denoted by $R_i(k) = \{j \in \Gamma(i) \mid k \in \Gamma(j)\}$. For example, in Fig. 1, $R_i(k) = \{j, m\}$. Note that, n is not in $R_i(k)$ since n does not know k . Every agent in $R_i(k)$ gives a recommendation to i .

Once the agent receives the recommendations, the evaluation process begins. Evaluation of the recommendations by agents varies according to their dispositions. We consider three types of dispositions:

(i) *Optimists*. An optimist agent i takes the maximum of the recommendations that it receives [4], that is,

$$t_{ik} = \max\{t_{jk} \mid j \in R_i(k)\}.$$

(ii) *Pessimists*. A pessimist agent i takes the minimum of the recommendations that it gets [4], namely,

$$t_{ik} = \min\{t_{jk} \mid j \in R_i(k)\}.$$

Optimist and pessimist agents are the opposites of each other. We consider a third type in between.

(iii) *Realists*. A realist agent i takes the average of the recommendations [4], that is,

$$t_{ik} = \frac{1}{|R_i(k)|} \sum_{j \in R_i(k)} t_{jk}.$$

In the literature, there are realist agents that use the mode or the median of the recommendations, too [15].

(iv) *Self-assured (SA)*. To compare our agents with the previous work [12], [13], we also consider agents that do not ask for recommendations. They use their memory only.

D. Perceived cooperation ratio

We modify perceived cooperation ratio given in Eq. 1. According to Eq. 1, if agent j plays with i once and cooperates, t_{ij} will be 1. Similarly, if j plays with i , say 10 times and cooperates in all of them, t_{ij} is still equal to 1. However, in real life, the trustworthiness of a person who was honest with us only once and that of a person who was honest with us many times are not the same [16]. With this idea, we propose a different approach to calculations of perceived cooperation ratio. *Perceived cooperation ratio* of j with respect to i is defined as

$$t_{ij} = \frac{c_{ji} + 1}{(c_{ji} + 1) + (d_{ji} + 1)}. \quad (2)$$

With our new formula, if j plays with i once and cooperates, t_{ij} becomes 0.66. On the other hand, if j plays with i 10 times and cooperates in all of them, t_{ij} becomes 0.91. We feel that Eq. 2 provides more realistic perception evaluation. Note that both Eq. 1 and Eq. 2 produces the same value of $t_{ij} = 0.5$ when $c_{ji} = d_{ji}$. Therefore, since the maximum or the minimum are still either larger or smaller than 0.5, optimist and pessimist agents are not affected from this new definition of perceived cooperation ratio. On the other hand, realist agents are affected since average will be different.

E. Metrics

We want to compare the performances of the cooperators and the defectors at the end of the game. To do that we define *average payoff* of all agents in a set A as

$$\overline{P_A} = \frac{1}{|A|} \sum_{i \in A} \text{payoff}(i).$$

where $\text{payoff}(i)$ denotes the accumulated payoffs by agent i at the end of the game.

We have two sets of agents. The set C of agents with cooperation probability $\rho > 0.5$ are considered cooperators. The rest of the agents are called defectors and denoted by D . Now, we can define *payoff ratio* as

$$\phi_C = \overline{P_C} / \overline{P_{C \cup D}}.$$

Note that we are interested in cases of $\phi_C > 1$, where cooperators have higher average payoff

than the average, that is, they perform better than defectors.

It is also possible that agents can misjudge the opponent although they collect the recommendations. The number of misjudgments of cooperators as defectors is denoted by η_{cd} . Similarly, η_{dc} is the number of misjudgments of defectors as cooperators. Then, the *accuracy ratio* of evaluations is defined as the ratio of the total number of correct judgments to the total number of recommendations received η_r , namely,

$$\delta = 1 - \frac{\eta_{cd} + \eta_{dc}}{\eta_r}.$$

IV. RESULTS

Given the model, it is possible to come up with many possible scenarios. In this paper, we investigate some of them. (i) We consider homogeneous agents, i.e., one type of cooperators with cooperation probability $\rho = 0.9$ playing against one type of defectors with $\rho = 0.1$. (ii) In our experiments, both cooperators and defectors use the same forget strategy such as forget cooperators first. (iii) Defectors are always *self-assured*, i.e. they do not ask for a recommendation.

In contrast to defectors, cooperators do ask for a recommendation. Therefore, for a given parameter set, an experiment actually consists of four different simulations where cooperators are either (i) optimistic or (ii) pessimistic or (iii) realist or (iv) self-assured playing against self-assured defectors. We will be using plots such as Fig. 3(a) to compare performances of different dispositions.

In each experiment, there are $N = 100$ agents, where 50 defectors play against to 50 cooperators, i.e., 50% cooperators. We terminate the experiments after $\tau \binom{N}{2}$ pairs invited to play as in the case of Cetin and Bingol's model [13], where $\tau = 30$. We use payoff matrix of $(S, P, R, T) = (0, 1, 3, 5)$. We report the average of 10 realizations.

We run experiments for various memory ratios and report our findings as a function of the memory ratio of μ .

A. Self-assured agents

First of all, we consider agents that do not get recommendations as in the case of Cetin and Bingol's model [12]. Fig. 2 agrees with the finding

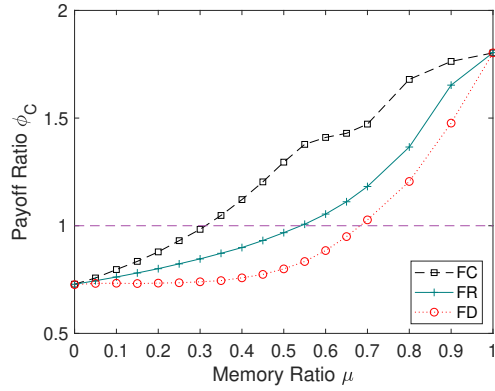


Fig. 2: Comparison of forget mechanisms for self-assured agents.

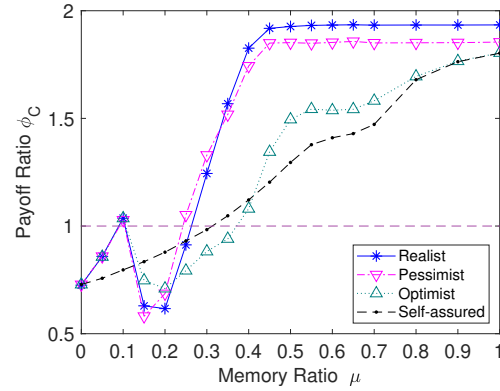
of Ref [12] that forgetting cooperators first (FC) is a better strategy compared to random forgetting (FR) or forgetting defectors first (FD). Even for FC, in order for cooperators to go above average payoff, i.e., $\phi_C = 1$, considerable memory ratio, $\mu = 0.3$, is required. Strategies FR and FD call for memory ratios more than 0.5 for $\phi_C > 1$.

B. Forget cooperators first

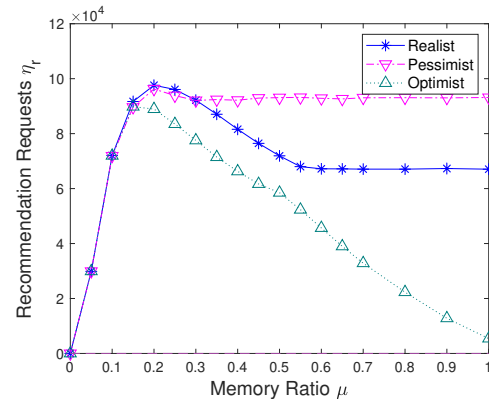
We investigate forgetting cooperators first strategy since it is better compared to FR and FD strategies according to Fig. 2 and Ref [12]. Co-operators of a disposition, such as optimist, play against self-assured defectors. Both cooperators and defectors use *forget cooperators first* (FC) strategy. That is, if there is no space left in the agent's memory, a cooperator in the memory is randomly selected and forgotten. If there is no cooperator left in the memory, then a randomly selected defector is forgotten.

1) *Performance*: As expected, Fig. 3(a) shows that self-assured agents increase their performance as memory ratio increases. Note that the self-assured curve in Fig. 3(a) and the FC curve in Fig. 2 are the same.

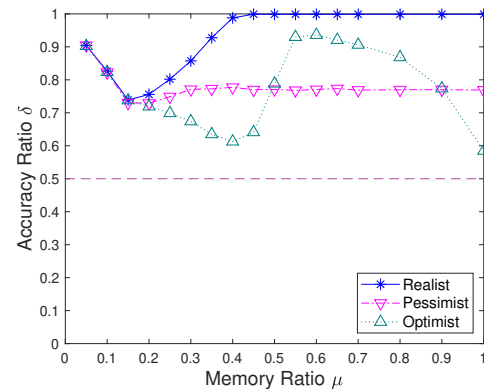
Interestingly, in Fig. 3(a), we observe unexpected fluctuations in the performances of the optimists, pessimists and realists around $\mu = 0.2$. Instead of increase, they decrease. Note that, although agents receive recommendations, their payoff ratio is worse than self-assured agents in this region.



(a) ϕ_C - Payoff ratios.



(b) n_r - Number of recommendation received.



(c) δ - Accuracy ratios of evaluations.

Fig. 3: Detailed analysis of forgetting cooperators first. (a) describes payoff ratios of agents; (b) describes the number of recommendation they receive; (c) describes the accuracy ratios of evaluations of agents

For the values of $\mu > 0.2$, performances start to increase again but different dispositions have different trends. Pessimists and realists are quick to recover and pass $\phi_C = 1$ threshold around $\mu = 0.25$. They reach their peak values on around $\mu = 0.45$ and stay there. Optimists have a different path. They cross $\phi_C = 1$ threshold around $\mu = 0.35$. Their performance steadily increases and takes its peak value close to $\mu = 1$.

2) *Number of requests*: In order to understand the strange behavior in performance, we collect data on recommendations.

There are recommendation requests, which receive no response. Then, agent has to play. If it receives at least one response, then it acts accordingly. Fig. 3(b) plots η_r , the number of recommendation requests that received at least one response. For all three dispositions, η_r increases as μ increases but it reaches to its maximum around $\mu = 0.2$. Then different dispositions present different behaviors. Pessimists keep the same level around $\phi_C = 9.5 \times 10^4$ for $\mu > 0.3$. For realists, the number of recommendation requests decreases till $\mu = 0.55$ and stays around $\phi_C = 6.5 \times 10^4$. For optimists, after its early peak value at $\mu = 0.15$, the number of recommendation requests decreases almost to none.

3) *Accuracy ration*: In Fig. 3(c), δ values of cooperators are shown. Since 50% of the population is cooperators, $\delta = 0.5$ threshold is marked with a dashed line. Note that the accuracies of all dispositions are above this line.

Initially, accuracy decreases as the memory of the agents' increases for all dispositions. Around $\mu = 0.15$, dispositions start to deviate. For realists, accuracy increases and reaches its maximum value of perfect accuracy, i.e. $\delta = 1$, around $\mu = 0.4$. The accuracy of pessimists increases slightly and stays just below $\delta = 0.8$ for $\mu > 0.3$. Behavior of optimists is the most difficult one to explain. It keeps decreasing to just above $\delta = 0.6$ as μ approaches to 0.4, then it has a sharp increase that reaches to above $\delta = 0.9$ at $\mu = 0.55$, and then has a smooth decrease back to $\delta = 0.6$.

V. DISCUSSION

Note that the way we set the experiment, cooperators are of one type of disposition only, say realists. Then, all cooperators are realist. That

is, any recommendations from cooperators are coming from realist cooperators. And they are evaluated by the agent which is also a realist. We discuss each disposition separately.

A. Optimists

Optimists, by their nature, have an optimistic view of life. One positive recommendation is good enough for them to play. Once they play with an agent, they record this firsthand experience in their memory. Because of this, they do not need to ask recommendations for the same agent again. This explains the steady decrease of recommendation requests in Fig. 3(b) for $0.2 < \mu < 1$. For $0.2 < \mu < 0.4$ region in Fig. 3(c), its accuracy keeps decreasing, too. That is, it requests less recommendation and makes bad judgements. Yet it knows enough defectors so that it can keep its payoff ratio increasing in this region..

B. Pessimists

Pessimists have the opposite strategy in playing. One single negative recommendation is enough for a pessimist not to play. If it does not play, then there will be no record of the opponent kept in its memory. If the same opponent is matched to play again, a pessimist has to ask recommendation once more. This explains the high number of queries of recommendation even at $\mu = 1$, where there is enough memory to keep the entire population.

C. Realists

Realists are in the middle ground of optimists and pessimists. They play more than pessimists but less than optimists. In the region of $0.2 < \mu < 0.6$ of Fig. 3(b), similar to optimists, they play with new agents. From that point on, their actions are similar to that of pessimists, that is, reject to play and keep asking the same agent over and over again.

D. Anomaly around $\mu = 0.2$

We further investigate the reasons for the surprising behavior in the region $0.15 < \mu < 0.25$. First, we ask whether they receive recommendations or not? By checking η_r values around $\mu = 0.2$, one can observe that the optimist,

pessimist and realist agents received sufficient recommendations. Then, we need to investigate how accurately they evaluate those recommendations. One observes that accuracy declines as μ go from 0.1 to 0.2, which can explain the drop in performance ϕ_C in the same region.

In this region around $\mu = 0.20$, memory is quite small. One can hold at most 25% of the population while 50% of it are defectors. Because of the forgetting strategy of FC, one keeps perceived defectors in the memory. Therefore any recommendation given would be towards no play. If a neighbor misjudges a cooperator as a defector, this misjudgment prevents others from play with that cooperator. This escalation prevents others from play with this cooperator and gain points.

E. Possible extensions

This work can be expanded in various ways. First, in our current model, the recommender agent gives its sincere perception as a recommendation. That is, even a defector agent provides its genuine opinion. This may not be the case in real life. Second, we considered $N = 100$ with 50% cooperators. One wants to try larger N values as well as different percentages. Third, we consider homogenous agents. One type of cooperators plays against one type of defectors. In real life, there are always mixtures of all kinds. In such heterogeneous environments are difficult to investigate but definitely much more realistic. Fourth, there are possibilities for defectors. Our defectors are self-assured. One may consider defectors that use recommendations. More than that, our defectors refuse to play with perceived defectors. Given a payoff matrix of $(S, P, R, T) = (0, 1, 3, 5)$, a realist defector would always choose to play since it has nothing to loose. Payoff matrices with negative entries as in the case of [13] is another possibility to investigate. Finally, any two agents play τ times on average. Different τ values should be investigated.

VI. CONCLUSION

We investigated the Iterated Prisoner's Dilemma game where agents get recommendations if they do not know the opponent. Although we expect better performance as memory capacity increases,

the performances of all dispositions drop around $\mu = 0.2$ region, in which agents that do not get any recommendations perform better. After that region performances recover. Realists have the best performance while the optimists have the worst.

In this work, we report, in detail, strong co-operators and strong defectors with $\rho = 0.9$ and $\rho = 0.1$, respectively. We also investigate mild cooperators and defectors such as $\rho = 0.75$ and $\rho = 0.25$, and we obtained similar results.

ACKNOWLEDGMENTS

We would like to thank Uzay Cetin, Mursel Tasgin, and Emre Aladag for constructive comments. This work was partially supported by the Turkish State Planning Organization (DPT) TAM Project (Grant Number: 2007K120610).

REFERENCES

- [1] H. Bingol, "Fame emerges as a result of small memory," *Physical Review E*, vol. 77, no. 3, p. 036118, 2008.
- [2] D. Susan and J. G. Holmes, "The dynamics of interpersonal trust: Resolving uncertainty in the face of risk," *Cooperation and Prosocial Behavior; Cambridge University Press: New York, NY, USA*, p. 190, 1991.
- [3] W. Sherchan, S. Nepal, and C. Paris, "A survey of trust in social networks," *ACM Computing Surveys (CSUR)*, vol. 45, no. 4, p. 47, 2013.
- [4] S. Marsh, "Optimism and pessimism in trust," in *Proceedings of the Ibero-American Conference on Artificial Intelligence (IBERAMIA'94)*, 1994.
- [5] R. Axelrod and W. D. Hamilton, "The evolution of cooperation," *Science*, vol. 211, no. 4489, pp. 1390–1396, 1981.
- [6] M. Leas, E. Dolson, R. Annis, J. Nahum, L. Grabowski, and C. Ofria, "The prisoners dilemma, memory, and the early evolution of intelligence," in *Proceedings of the Artificial Life Conference 2016 13*. MIT Press, 2016, pp. 408–415.
- [7] M. A. Nowak, "Five rules for the evolution of cooperation," *science*, vol. 314, no. 5805, pp. 1560–1563, 2006.
- [8] C. Chu, J. Liu, C. Shen, J. Jin, and L. Shi, "Win-stay-lose-learn promotes cooperation in the prisoner's dilemma game with voluntary participation," *PloS one*, vol. 12, no. 2, p. e0171680, 2017.
- [9] H. Yamamoto, I. Okada, T. Taguchi, and M. Muto, "Effect of voluntary participation on an alternating and a simultaneous prisoner's dilemma," *Physical Review E*, vol. 100, no. 3, p. 032304, 2019.
- [10] M. A. Nowak and K. Sigmund, "The dynamics of indirect reciprocity," *Journal of theoretical Biology*, vol. 194, no. 4, pp. 561–574, 1998.
- [11] —, "Evolution of indirect reciprocity by image scoring," *Nature*, vol. 393, no. 6685, p. 573, 1998.
- [12] U. Cetin and H. O. Bingol, "Iterated prisoners dilemma with limited attention," *Condensed Matter Physics*, vol. 17, no. 3, p. 33001, 2014.

- [13] —, “The dose of the threat makes the resistance for cooperation,” *Advances in Complex Systems*, vol. 19, no. 08, p. 1650015, 2016.
- [14] E. A. Stanley, D. Ashlock, and M. D. Smucker, “Iterated prisoner’s dilemma with choice and refusal of partners: Evolutionary results,” in *Proceedings of the Third European Conference on Advances in Artificial Life*. London, UK, UK: Springer-Verlag, 1995, pp. 490–502.
- [15] S. P. Marsh, “Formalising trust as a computational concept,” Department of Computing Science and Mathematics University of Stirling, Tech. Rep., 1994.
- [16] R. Kiefhaber, R. Jahr, N. Msadek, and T. Ungerer, “Ranking of direct trust, confidence, and reputation in an abstract system with unreliable components,” in *2013 IEEE 10th International Conference on Ubiquitous Intelligence and Computing and 2013 IEEE 10th International Conference on Autonomic and Trusted Computing*. IEEE, 2013, pp. 388–395.