

The Success and Failure of Tag-Mediated Evolution of Cooperation

Austin McDonald and Sandip Sen

Mathematics and Computer Science Department
University of Tulsa
{austin, sandip}@utulsa.edu

Abstract. Use of tags to limit partner selection for playing has been shown to produce stable cooperation in agent populations playing the Prisoner’s Dilemma game. There is, however, a lack of understanding of how and why tags facilitate such cooperation. We start with an empirical investigation that identifies the key dynamics that result in sustainable cooperation in PD. Sufficiently long tags are needed to achieve this effect. A theoretical analysis shows that multiple simulation parameters including tag length, mutation rate and population size will have significant effect on sustaining cooperation. Experiments partially validate these observations. Additionally, we claim that tags only promote mimicking and not coordinated behavior in general, i.e., tags can promote cooperation only if cooperation requires identical actions from all group members. We illustrate the failure of the tag model to sustain cooperation by experimenting with domains where agents need to take complementary actions to maximize payoff.

1 Introduction

Learning and reasoning in single or multistage games have been an active area of research in multiagent systems [1–7]. Most of this research has concentrated on simultaneous move games with solution concepts like Nash equilibria [8, 9]. Nash Equilibrium, however, does not guarantee that agents will obtain the best possible payoffs, i.e., Nash Equilibrium does not ensure Pareto-optimal solutions. Some non-Nash Equilibrium action combinations may yield better payoffs for both agents, which may be reached if the agents look ahead to future iterations of the game while selecting actions [10].

That Nash Equilibria may not be the preferred outcome is particularly evident in the widely-studied Prisoner’s Dilemma (PD) game (see Figure 1). In this game, the only Nash Equilibria is the strategy profile (D,D) which is also the only non-Pareto-optimal outcome! The (D,D) strategy profile is dominated by the (C,C) strategy profile. Unfortunately, in a single-shot PD game, rational play will produce the Nash Equilibrium strategy profile. In repeated or iterated play, however, learning approaches can produce higher payoff by choosing the (C,C) strategy profile. Numerous researchers in game theory and in multiagent systems have attempted various mechanisms to induce cooperation in iterated PDs [11, 12, 7, 13–15].

We are particularly interested in recent work using tags in a population of interacting players (agents) [16, 17]. Tags have been proposed by John Holland as a primitive means of communication that can aid in the evolution of a group [18–20]. A tag is

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

Fig. 1. Utilities to players in a two-player Prisoner’s Dilemma game. Constraints on the utility values are $T > R > P > S$ and $2R > T + S > 2P$.

a label associated with an individual, often encoded as a bit string or a real number, that specifies the set of other population members with which this member is allowed to interact. Interaction between two individuals are allowed if their tags match, where matching often means that the tags are identical or that their difference, as measured by some distance function, is below a specified threshold. Tags have also been used by other researchers to promote cooperation in variations of PDs [21, 22]. Whereas these papers provided a reasonable high-level explanation of how the use of tags promotes cooperation, a detailed analysis that clearly explains the fundamental subtleties of the interactions in the population was missing. As a result, design of tag systems was based on trial and error and did not explain why certain parameter choices for such systems succeeded in inducing cooperation whereas others did not.

Another key observation was that all the domains to which tags have been applied so far contained a narrow characterization of cooperation as imitation of behavior. In multiagent domains in general, cooperation requires a richer, divergent collection of behaviors. It was unclear from the current state of knowledge whether tags can support cooperation in a broad spectrum of multiagent problems.

In this paper, we plan to carefully characterize the detailed interactions of agents using tags. The goal is to identify domains where tags would be useful for promoting cooperation and to develop a methodology for choosing parameter values in the tag framework that will actually facilitate the evolution of cooperation in such domains.

2 Related Work

The usage of tags to bias interactions on iterated Prisoner’s Dilemma has been suggested by [23, 19, 18]. Riolo performed the pioneering experiments in [21]. Riolo’s agents were modeled as a stochastic strategy, based on Tit-For-Tat, combined with a real-valued tag and a real-valued bias, both on the interval $[0, 1]$. Agents then attempt to pair up, where the difference between the agents’ tags is less than each agents’ bias. If no suitable pairing can be found within a small number of tries, the agent simply chooses a partner at random. When each agent has an identical, fixed bias, Riolo’s model results in increased performance for the society. However, when each agent’s bias is allowed to evolve, behavior varies drastically according to initial conditions and the results are less clear.

Hales and Edmonds [17] used a different model where the population consists of a collection of agents represented as a binary string of $l + 1$ bits. The first bit represents a pure strategy (always cooperate or always defect), while the remaining l bits are the tag. In each population generation, every agent plays a PD game against one other agent

with an identical tag. If an agent has a unique tag in the population, it plays against a randomly selected opponent. The next generation is formed via fitness proportionate reproduction where the fitness of an agent is the payoff received in this round of play. Mutation is then applied to each bit. This process is described in Algorithm 1.

Algorithm 1 Hales and Edmond’s model of population evolution with tags.

```

for some number of generations do
  for each agent  $a$  in the population do
    Select a game partner agent  $b$  with the same tag (if possible)
    Agent  $a$  and  $b$  invoke their strategies and  $a$  gets corresponding payoff.
  end for
  Reproduce agents in proportion to their average payoff (with some low level of mutation)
end for

```

3 Tag-based Model for Population Evolution

Our model is characterized as follows: Each agent in a population of size N is represented by two numbers: an integer s representing its strategy $0 \leq s \leq$ (number of valid strategies), and a real number t representing its tag $0 \leq t \leq 1$. The real number representation of the tag provides for an infinite tag space. Additionally, we divide the interval $[0, 1]$ into subintervals of length δ called *tag groups*.

The population is evolved over a number of generations. Each generation, every agent chooses another agent in the same tag group and plays a round of the game with the other agent. If an agent is the only member of a tag group, it randomly chooses another agent from the population. The first agent receives the payoff from that round. This process is repeated for all agents. Thus, an agent may interact with many agents per generation, but receives only one payoff per generation. This prevents singleton agents from interfering with other groups, while still giving them a chance to play, and hence survive.

The next generation is selected by a payoff proportionate reproduction scheme. For each selected agent the mutation operator, which replaces the current value of a parameter with a new number randomly chosen from the range for that parameter, is applied to the strategy with probability μ_S and to the tag with probability μ_T . The new generation is then evolved as described above.

The goal of this framework is to enable consistent evaluation of the performance of agent societies on a variety of games. Our proposed setup can be used to approximate a variety of existing tag models. It is similar to Riolo’s model using a global fixed bias, with two exceptions: we use a pure strategy, and being in the same tag group is transitive in our model (but such is not necessarily the case in Riolo’s). To compare to Hales’ and Edmonds’ model, we observe that there are 2^l distinct tag groups possible when using a tag with l bits. So we use $\frac{1}{\delta} = 2^l$, which lets us choose a corresponding δ for any given

l . Additionally, in Hales' and Edmonds' original model¹, the mutation rate of the tag was a function of tag length, which is not the case here.

4 Tags in Prisoner's Dilemma and related games

We recount Hales' [17] explanation of how tags help promote cooperation in the PD game. A homogeneous group of cooperators will prosper and grow. When such a group is invaded, via mutation of the strategy or tag, by a defector, the defector will prosper, resulting in imitators in the next generations. Over time, the group will fill with defectors, resulting in worsening performance and eventual extinction. Thus defectors, even if formed by chance, will not live long, and hence a majority of the population will be cooperators.

Our additional observation is that individuals with unique tags, i.e., singletons, can prosper if the randomly chosen individual they interact with is cooperative in nature. If the singleton is cooperative, then it will perform well, leading to perhaps more agents copying its tag bits and strategy and thus the group expands as a group of cooperators. If the singleton is a defector, when others copy its strategy and tag, all agents in this group will be defectors, they will perform poorly in the following generation and the group will die out.

We were puzzled, however, by the observation by Hales that sufficiently long tags were required to sustain cooperation in the PD game. For example, Hales observed that while 32 bit tags were able to sustain cooperation, 8 bit tags were not for a population of 100 agents! Note that this translates into the requirement of huge tag spaces, e.g., 32-bit tags produce a tag space of 2^{32} ! The explanation of how tags facilitate cooperation has no apparent requirement for such massive tag spaces. We also did not find any theoretical justification in Hales's work on characterizing the effect of tag length on promoting cooperation. This led us to believe that the current understanding of the working of the tag mechanism is incomplete and we need to improve on it in order to design working tag-based systems for arbitrary populations playing the iterated PD game and for other applications.

We are interested in investigating whether longer tags are required solely for the associated increase in mutation rate, or whether there were other beneficial effects. Additionally, we sought to validate the claim that tags are useful only in cases where mimicry is a desirable strategy.

To investigate tag length, we chose a game that tags were known to perform well on: the Prisoner's Dilemma (Figure 1). We chose $T=1.9$, $R=1.0$, $P=.002$, and $S=.001$ in order to accentuate the differences in the outcomes. Using closer values yields results that are similar to those presented here. Unless otherwise stated, all experiments use $\mu_s = \mu_t = .1$. Since we are primarily interested in comparisons to Hales' and Edmonds' work, we will specify δ in terms of bits of tag length, l . The average payoffs per agent over time in a set of typical Prisoner's Dilemma runs are presented in Figure 2. The left plot is equivalent to not using tags, and is presented to demonstrate that tags do produce improvement. We see that for shorter tags, the payoffs are clustered around a lower average than for progressively higher tag lengths.

¹ This was changed in a later paper [24].

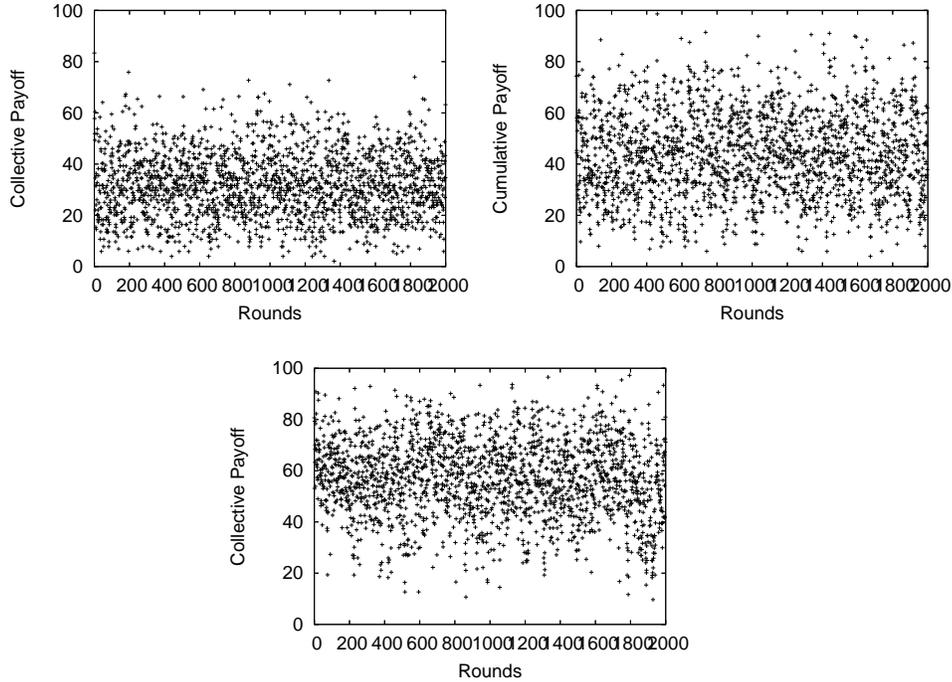


Fig. 2. Average payoffs for $l=0$ (left), 4 (right), and 32 (bottom) for the Prisoner's Dilemma game.

The population characteristic that we found most distinguished runs with long versus short tags was the number of populated groups. We present the corresponding plot for the number of groups over a run in Figure 3. The left plot uses $l=8, \mu_T=.001$ and the right plot uses $l=32, \mu_T=.1$. The figures clearly denote significant differences in the population features with larger number of groups and smaller average group sizes in the case of longer tags and higher mutation rates.

This observation is key to the following conjecture: *Cooperation is sustained in a population of tag-based players playing PD and evolved based on fitness if sufficient number of cooperative groups are created via mutation.* The conjecture is supported by the following reasoning. If there are too few cooperative groups, invasion by defectors will destroy them. If there are sufficiently many cooperative groups, destroying a few will still leave the possibility of other groups spawning more new cooperative groups via mutation. The basic argument is that there has to be enough groups such that the rate of destruction via invasion by defectors is less than the formation of new groups by mutation.

To further investigate this conjecture, we now attempt a partial theoretical characterization of this phenomena. We first calculated the expected number of new groups formed each generation, M .

$$M = N(1 - \delta g)\mu_T$$

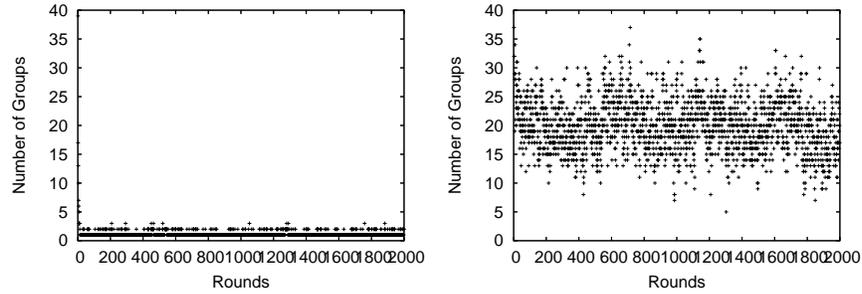


Fig. 3. Number of populated groups in Prisoner's Dilemma for $l = 8$, $\mu_T = .001$ (left) and $l = 32$, $\mu_T = .1$ (right).

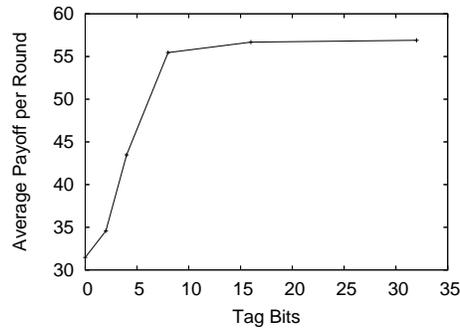


Fig. 4. Average payoffs per round with decreasing δ (increasing tag length).

where g is the current number of populated groups and N is the population. This is only an approximation; after a new group is formed, the value of g changes. However, when $N \ll \frac{1}{\delta}$, this approximation is valid. From this equation, consider $P = M/N$, the fraction of agents in the population that formed new groups. Higher values for P can be thought of as having two complementary effects: the likelihood of an agent forming a new group is increased, and the likelihood of an agent joining an existing group is decreased. This directly increases the rate of formation of new cooperative groups (proportional to the aggregate payoffs of cooperative agents) while decreasing the rate that groups are invaded by defectors from outside. However, P cannot be allowed to become too high; this would result in too many agents forming new groups and not staying behind to achieve the cooperative payoff.

So by varying P , we should enhance the payoffs obtained by a population. However, we cannot directly lower g , even if doing so were to be beneficial. So instead we focus on δ and μ_T . By simply lowering δ to zero, we can ensure that any mutation will result in a new tag group being formed (note that this would correspond to an infinite tag length). Then, we will have $P = \mu_T$. Note that we have not lost generality, since all values of P we could achieve with any given δ are still achievable with μ_T . We present the effect of progressively lowering δ in Figure 4.

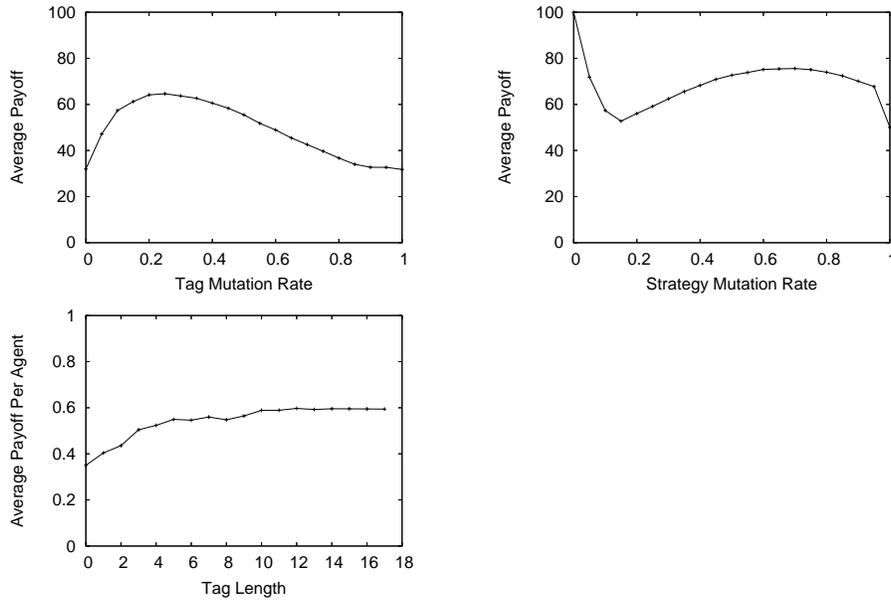


Fig. 5. Average payoffs as we increase μ_T (top left) μ_S (top right) and N (bottom).

The above characterization explains how smaller δ (or equivalently, longer tags) can increase the performance of a population on Iterated Prisoner's Dilemma (up to a certain bound) despite a fixed mutation rate. According to our theory, then, when $\delta = 0$ there should be a μ_T with which we can achieve the best performance possible from the system. When $\delta = 0$, the only way agents can be in the same group is if they have identical real-valued tags. Since the chances of an agent randomly mutating to an existing tag are infinitely small, agents with identical tags must have been reproduced from the same parent.² So, it is likely that they will share common strategy bits. This behavior leads us to speculate that tag systems are merely promoting mimicry, rather than cooperation.

Having decided on a fixed $\delta = 0$, we now investigate the effects of altering μ_T and μ_S . (Figure 5). From the figure, we see that there exists an optimal value for μ_T near $\mu_T = .2$. Higher values result in groups not remaining together long enough for cooperation to occur; lower values result in fewer new groups being formed, and thus a higher proportion being lost due to defection. With regards to μ_S , we conjecture that higher values are causing agents to switch strategies too quickly to be able to cooperate, while lower values aren't producing enough cooperators.

² This is only theoretically true. In practice, machines have a finite precision for representing real numbers. However, this precision is typically large enough that we can consider it to be infinite.

5 Effect of Tags Where Cooperation \neq Mimicry

Several authors have suggested other applications of tags [25, 16, 22]. However, we observed that in all of the applications mentioned mimicry is an effective strategy for increasing performance. Though mimicking behavior can produce cooperation in PD and other domains, for a large gamut of multiagent interactions complementary, rather than identical behavior is required for cooperation. For tags to be used as an effective facilitator for promoting cooperation in general, it is imperative that we better understand their role in aiding cooperation through complementary behavior. Our own experiments (see above) suggest that tags may only be helpful in games where cooperative actions are identical.

To investigate the performance of tag systems on a game where complementary actions are required to maximize payoff, we look at the performance of a tagged system on a game where mimicry is useless: the Anti-Coordination game. In this game, cooperation occurs when players choose complementary actions; note that this game is not a dilemma (ie, one player cannot succeed at the other player's expense). We expect that tags can improve the performance of a society on this game by increasing the likelihood of pairing a 1 with a 0 from the expected equilibrium strategy without tags, which is an even spread across the strategies (half 1's, half 0's). Next we present similar results for the anti-coordination game (Figure 6).

	0	1
0	L, L	H, H
1	H, H	L, L

Fig. 6. Utilities to players in Anti-Coordination. Constraints on the utility values are $H > L$.

Figure 7 shows the average payoffs of a society over time with varying tag lengths. We see that changing l has no noticeable effect on the population. We show the effects of changing l over time in Figure 8, where each point is an average of 5 runs. We also present the average performance of the population as we alter the strategy mutation rate μ_S while holding fixed $\delta = 0$. The tagging mechanism has no effect on the society playing this game.

To explain why tags fail to increase performance on this game, assume we begin with an equally distributed population of 0's and 1's. Without tags, this is the optimal distribution. Any increase in 0's or 1's will be detrimental to the society. For example, say the number of 0's increases. Then the average 0 will receive a lower payoff than the average 1, and so the number of 0's will decrease and the number of 1's will increase, pushing the system back to equilibrium.

In order to improve the performance of this system, we need to restrict the agents such that they only play with other agents who have opposite tags. As we discovered in our experiments with the Prisoner's Dilemma, the chance of an agent joining a tag group from outside is vanishingly small (and gets smaller as δ decreases). So in all likelihood any agent with the same tag as another agent is a copy of that agent (either

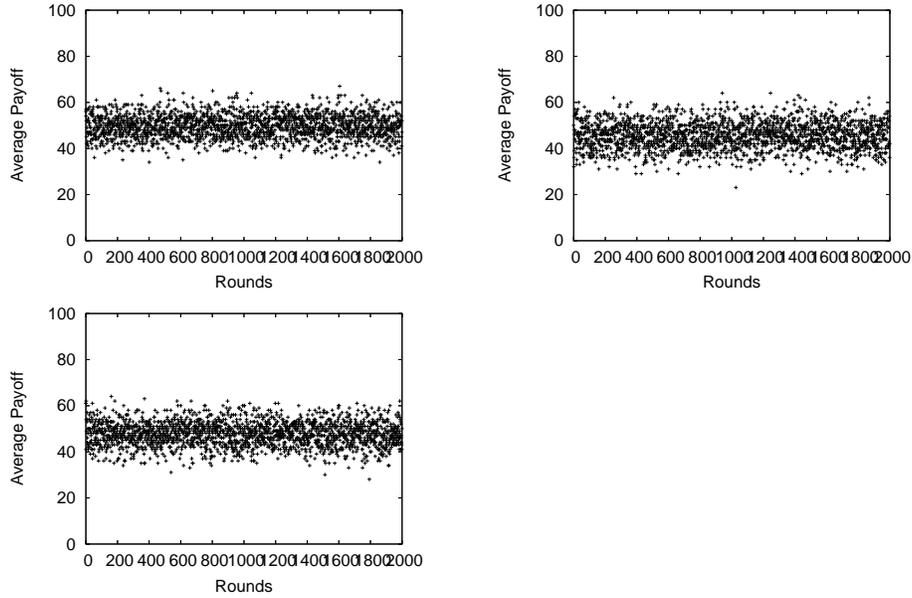


Fig. 7. Average payoffs for $l = 0$ (left), 8 (middle), and 32 (right) for the Anti-Coordination game.

directly or as a copy of a copy). So they will tend to have the same tag bit, and will both get low payoffs.

By this mechanism, the anti-coordination game actually discourages growth of homogeneous strategy groups, which totally negates the function of tags.

6 Conclusion

In this paper we show how and when tag-mediated agent interaction can promote and sustain cooperative behavior. In particular, we analyze how size of the tag space, mutation rate, and population size effect the evolution of cooperation in populations repeatedly playing the Prisoner's Dilemma game. Additionally, we have identified the function tags perform on influencing such systems: they promote mimicry, and not necessarily cooperation. We demonstrated this by studying the performance of a tagged society playing the Anti-Coordination Game. Our analysis and observations suggest that existing tag models primarily promote cooperative behavior in games where mimicry is the best strategy. While this means that tags can improve group performance in coordination games and games like the PD, they are inadequate for sustaining cooperation in general multiagent situations where effective cooperation requires complementary, and not identical, behaviors by the cooperators.

Acknowledgments: This work has been supported in part by an NSF award IIS-0209208.

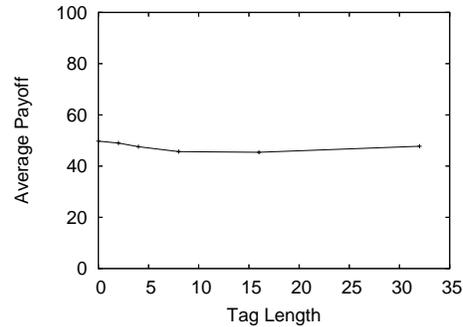


Fig. 8. Average payoffs as tag length increases.

References

1. Banerjee, B., Sen, S., Peng, J.: Fast concurrent reinforcement learners. In: Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence. (2001) 825–830
2. Bowling, M., Veloso, M.: Rational and convergent learning in stochastic games. In: Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence. (2001) 1021–1026
3. Claus, C., Boutilier, C.: The dynamics of reinforcement learning in cooperative multiagent systems. In: Proceedings of the Fifteenth National Conference on Artificial Intelligence, Menlo Park, CA, AAAI Press/MIT Press (1998) 746–752
4. Hu, J., Wellman, M.P.: Multiagent reinforcement learning: Theoretical framework and an algorithm. In Shavlik, J., ed.: Proceedings of the Fifteenth International Conference on Machine Learning, San Francisco, CA, Morgan Kaufmann (1998) 242–250
5. Littman, M.L.: Markov games as a framework for multi-agent reinforcement learning. In: Proceedings of the Eleventh International Conference on Machine Learning, San Mateo, CA, Morgan Kaufmann (1994) 157–163
6. Littman, M.L.: Friend-or-foe q-learning in general-sum games. In: Proceedings of the Eighteenth International Conference on Machine Learning, San Francisco: CA, Morgan Kaufmann (2001) 322–328
7. Littman, M.L., Stone, P.: Implicit negotiation in repeated games. In: Intelligent Agents VIII: AGENT THEORIES, ARCHITECTURE, AND LANGUAGES. (2001) 393–404
8. Myerson, R.B.: Game Theory: Analysis of Conflict. Harvard University Press (1991)
9. Nash, J.F.: Non-cooperative games. *Annals of Mathematics* **54** (1951) 286 – 295
10. Brams, S.J.: Theory of Moves. Cambridge University Press, Cambridge: UK (1994)
11. Axelrod, R.: The Evolution of Cooperation. Basic Books (1984)
12. Dugatkin, L.A.: The evolution of cooperation: four paths to the evolution and maintenance of cooperative behavior. *BioScience* **47** (1997) 355–361
13. Sigmund, K., Nowak, M.A.: The alternating prisoner’s dilemma. *Journal of Theoretical Biology* **38** (1994) 262–275
14. Stimpson, J.L., Goodrich, M.A., Walters, L.C.: Satisficing and learning cooperation in the prisoner’s dilemma. In: Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence. (2001) 535–540
15. Trivers, R.: The evolution of reciprocal altruism. *Quarterly Review of Biology* **46** (1972) 35–57

16. Hales, D., Edmonds, B.: Can tags build working systems? from mabs to esoa. In: Engineering Self-Organising Systems. Lecture Notes in AI-2977, Springer Verlag (2003) 186–194
17. Hales, D., Edmonds, B.: Evolving social rationality for mas using "tags". In: Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems, Melbourne, Australia, ACM Press (2003) 497–503
18. Holland, J.H., Holyoak, K., Nisbett, R., Thagard, P.: Induction: Processes of Inferences, Learning, and Discovery. MIT Press, Cambridge, MA (1986)
19. Holland, J.: The effect of labels (tags) on social interactions. Technical Report Working Paper 93-10-064, Santa Fe Institute (1993)
20. Holland, J.: Exploring the evolution of complexity in signalling networks. *Complexity* **7** (2001) 34–45
21. Riolo, R.: The effects and evolution of tag-mediated selection of partners in populations playing the Iterated Prisoner's Dilemma. In: Proceedings of the Seventh International Conference on Genetic Algorithms, Morgan Kaufmann Publishers, Inc. (1997) 378–385
22. Riolo, R., Cohen, M.D., Axelrod, R.: Cooperation without reciprocity. *Nature* **414** (2001) 441–443
23. Allison, P.D.: The cultural evolution of beneficent norms. *Social Forces* **71** (1992) 279–301
24. Hales, D., Edmonds, B.: Change your tags fast! - a necessary condition for cooperation? In: Proceedings of the Joint Workshop on Multi-Agent and Multi-Agent-Based Simulation. (2004)
25. Hales, D.: Self-organising, open and cooperative p2p societies - from tags to networks. In: Proceedings of the 2nd Workshop on Engineering Self-Organising Applications. (2004)