Analyzing the effects of tags on promoting cooperation in Prisoner's Dilemma

Austin McDonald and **Sandip Sen**
Department of Mathematics & Computer Science
University of Tulsa
austin-mcdonald@utulsa.edu, sandip@utulsa.edu

Abstract

In some two-player games, e.g., the Prisoner's Dilemma(PD), myopic decisions can produce poor performance both for the individual and the agent collection (society). When such games are played repeatedly between agents in a society, auxiliary mechanisms can be used to mitigate such lack of coordination. By biasing which other agents a particular agent can play, we can promote social coordination. Use of tags to limit partner selection for playing has been shown to produce stable cooperation in agent populations playing PD. There is, however, a lack of understanding of why sufficiently long tags are needed to achieve this effect. We empirically characterize the population features produced by longer tags that enable sustained cooperation. A theoretical analysis shows that similar effects can be obtained by increasing mutation rate and population size. Experiments partially validate these observations. We also predict that such increases may ultimately be detrimental at larger values.

Introduction

Learning and reasoning in single or multistage games have been an active area of research in multiagent systems (Banerjee, Sen, & Peng 2001; Bowling & Veloso 2001; Claus & Boutilier 1998; Hu & Wellman 1998; Littman 1994; 2001; Littman & Stone 2001). Most of this research has concentrated on simultaneous move games with solution concepts like Nash equilibria (Myerson 1991; Nash 1951). Nash Equilibria, however, do not guarantee that agents will obtain the best possible payoffs, i.e., Nash Equilibria do not ensure Pareto-optimal solutions. Some non-Nash Equilibria 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 (Brams 1994).

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 produce cooperation in iterated PDs (Andreoni & Varian 1999; Axelrod 1984; Dugatkin 1997; Ferriere & Michod 1996; Hales & Edmonds 2003b; 2003a; Ito & Yano 1995; Littman & Stone 2001; Posch 1999; Sigmund & Nowak 1994; Stimpson, Goodrich, & Walters 2001; Trivers 1972).

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

Figure 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.

We are particularly interested in recent work using tag bits in a population of interacting players (agents) (Hales & Edmonds 2003a; 2003b). Tags have been proposed by John Holland as a primitive means of communication that can aid in the evolution of a group (Holland *et al.* 1986; Holland 1993; 2001). Tags have also been used by other researchers to promote cooperation in variations of PDs (Riolo 1997; Riolo, Cohen, & Axelrod 2001).

Now, we explain how tags are used to influence interaction between agents. The population consists of a collection of agents, where each agent has associated with a binary tag string of l bits. In each generation of the population, each agent plays a PD game against another agent with an identical tag. If an agent has an unique tag in the population, it plays against a randomly selected opponent. A fitness proportionate reproduction mechanism is used to select the next generation where fitness of an agent is the payoff received in this round of play. This process is described in algorithm 1.

Hales (Hales & Edmonds 2003b) provides an explanation of how tags help promote cooperation. Here we summarize his argument. A group composed of cooperators will prosper and grow. When such a group is invaded by a defector, the defector will prosper, resulting in imitators in the next generations, and over time the group will fill with defectors, resulting in worsening performance and eventually extinc-

Algorithm 1 Outline algorithm for simulation of model M2

```
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 cor-
    responding payoff.
  end for
  Reproduce agents in proportion to their average payoff
  (with some, low, level of mutation)
end for
```

tion.

Our additional observation is that individuals with unique tags, 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.

The above extended explanation appears valid in our experiments. We were puzzled, however, by the observation by Hales that sufficiently long tags were required to sustain cooperation in the PD game. The above explanation has no critical requirement for lengthy tags. 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 story of tags, as described, is incomplete and we need to improve our understanding to be able to design working tag-based systems for arbitrary populations playing iterated PD game and for other applications.

This paper is our initial attempt in explaining the need for sufficiently long tag bits for promoting cooperation in a population playing the PD game. In the process we also discuss the effects of population size and mutation rate on promoting and sustaining cooperation.

Experimental Setup

An array of bits, of size $l + 1$, is used to represent each agent in a society of population N. Unless otherwise stated, our experiments used a population size of 100. The first l bits of the array are used as the *tag*, while the remaining bit is used as the strategy. Thus, for the PD, each agent uses a pure strategy of cooperation or defection.

The population is evolved over a number of generations. Each generation, every agent chooses another agent with the same tag as its own and plays a round of PD with the other agent. If no other agents with the same tag can be found, it simply chooses an agent at random from the population. The first agent receives the payoff from that round. This process is performed for all agents; thus, an agent may be on the non-payoff end of many different PD rounds, but only receives one payoff per generation. This prevents agents outside of a group from interfering with the group, yet still gives them a chance to survive.

After all agents have received a payoff, a weighted roulette wheel with replacement is used to select agents for the next generation. Mutation is applied to each bit of every agent with probability μ. Mutation results in flipping the bit. Unless otherwise specified, we use a μ value of .001. Then, the new generation is evolved as described above.

Our primary aim in the experimental process was to discern what the difference was in the population dynamics with larger versus smaller tag lengths. We found that the features we will discuss were easily comparable when using scenarios of 8 tag bits vs 32 tag bits. For example, when using 8 tag bits, the population often demonstrated sharp phase changes from mostly cooperator to mostly defectors and vice versa, whereas using 32 tag bits produced steady cooperation across runs. We now explore features of the evolution of population for these two different scenarios.

Results and Analysis

We first present the payoffs obtained by all agents in different generations of a run with 8 and 32 tag bits in Figure 2. These figures correspond closely with the number of cooperators in the population in different generations of a run for those tag lengths (see Figure 3).

A concept we found useful in analyzing the results was the idea of a "group." A group is a defined as a set of agents with identical tag bits. A "homogenous" group is one where all agents in the group have identical strategies, and these are further divided into "cooperating" groups and "defecting" groups. So, in a given experiment, there are 2^l groups with populations varying from 0 to n. A group composition changes during reproduction by four mechanisms: (1) agents joining or leaving the group due to tag mutation; (2) agents joining or leaving as a result of fitness-proportionate reproduction; (3) mutation of the strategy bit; and (4) a group may change from homogenous to heterogenous by an agent of differing strategy bit joining the group as (1) or (2) above; likewise, a heterogenous group may become homogenous if all of the differing agents leave. Obviously, larger tag sizes increase the likelihood of (1), (2), and (4) above, as does a higher tag mutation rate. An increase in strategy mutation rate only affects (3) above.

The population characteristics that we found most distinguished runs with long versus short tags was the number of groups and the number of agents per group. These numbers also followed the phase transitions from cooperation to defection as demonstrated with 8 tag bits. We present the corresponding plots for the number of groups over a run in Figure 4 and the average group sizes over a run in Figure 5. 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.

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.* The conjecture is supported by the following reasoning. If there are too few cooperative group, 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

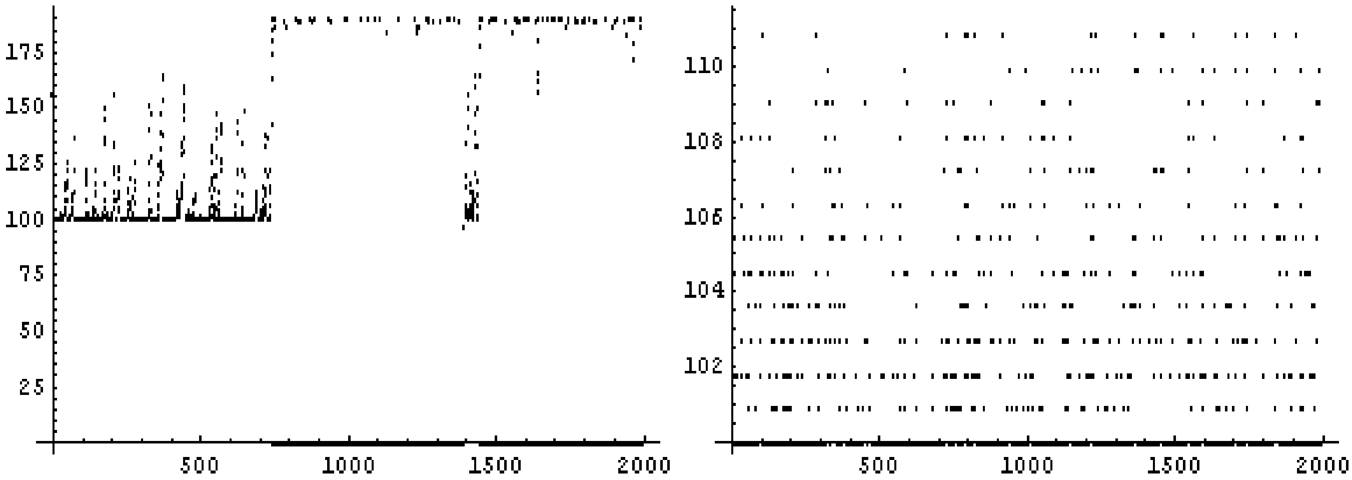

Figure 2: Payoffs over a typical run in experiments with 8 (left) and 32 (right) tag bits.

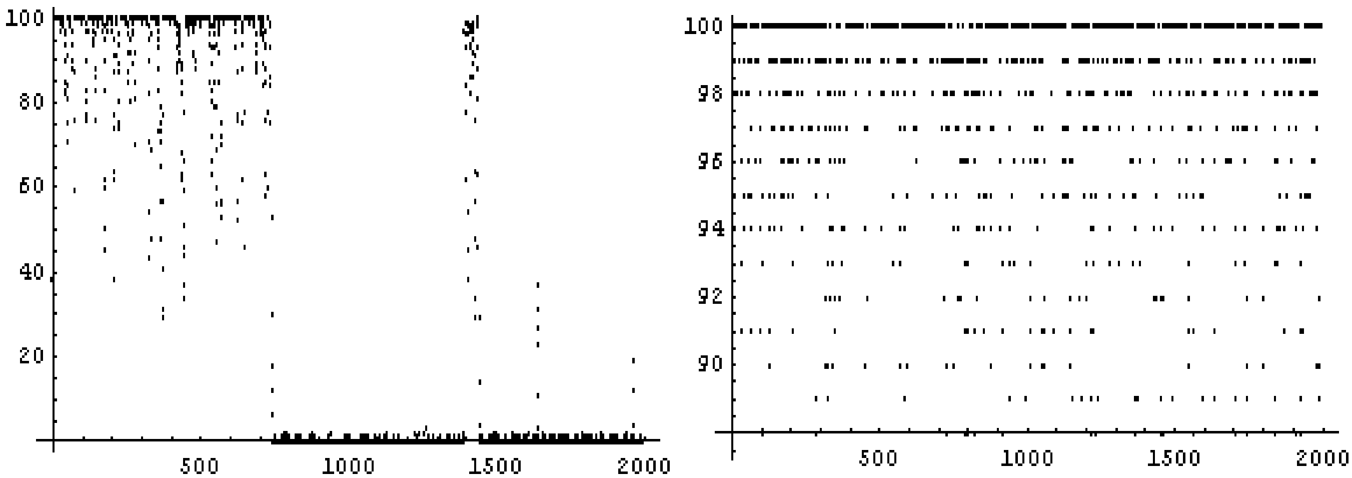

Figure 3: Number of cooperators in the population over a typical run in experiments with 8 (left) and 32 (right) tag bits.

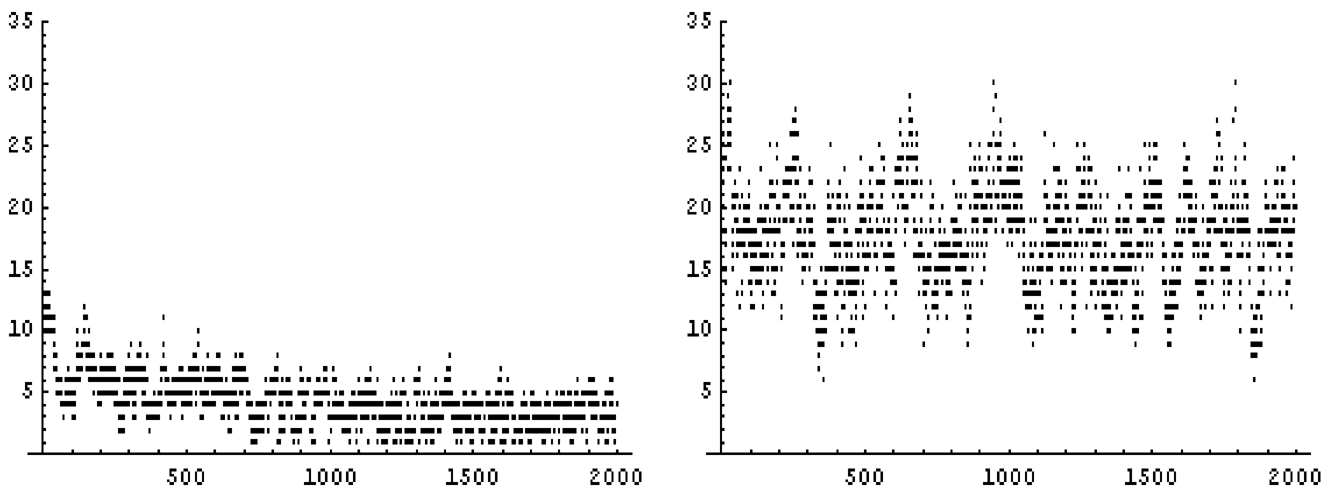

Figure 4: Number of groups over a run in experiments with 8 (left) and 32 (right) tag bits.

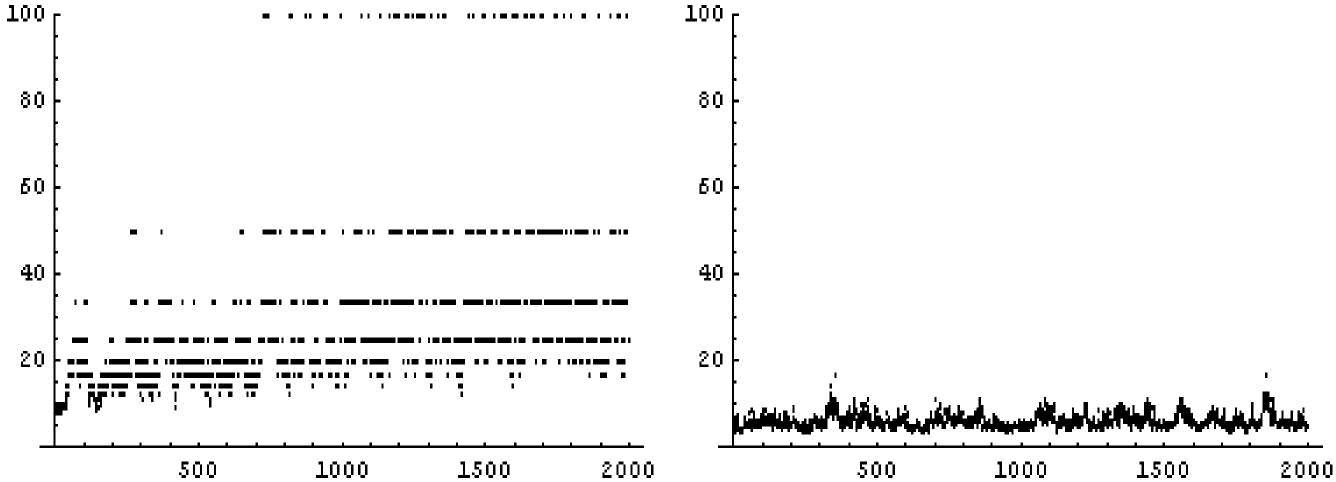

Figure 5: Average group size over a run in experiments with 8 (left) and 32 (right) tag bits.

more new cooperative groups via mutation. The conjecture is not strong enough in the sense of having a precise threshold of the number of groups beyond which cooperation will be self-sustaining. 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 probability that a given individual will mutate to a singleton to give rise to a new group. This probability is given by

$$\frac{2^l - g}{2^l}(1 - (1 - \mu)^l),$$

where g is the current number of groups. To simplify the next steps, we assume that the number of tag bits is large enough so that the size of the tag space is much larger than the population size, and hence much larger than the number of groups in the population. With this assumption (which may be violated in the case of large populations with small tag lengths), we have the expected number of singletons formed in a generation to be

$$M = N(1 - (1 - \mu)^l) \approx N\mu l,$$

where the last approximation is valid when $\mu l \ll 1$.

The above calculation also assumes that all the mutations produce singletons. This is not true for finite size tags, though is more likely with longer tags. The actual calculation involves a function $f(2^l - g, \lfloor M \rfloor)$ using a multinomial distribution without an easy to use closed form solution:

$$f(n,k) = \sum_{i=1}^{k} i \frac{\binom{n}{i} \sum_{e_1+\cdots+e_i=k, \forall i, e_i \geq 1} \frac{k!}{e_1! \cdots e_i!}}{n^k}.$$

We will use the simplified expression in the previous paragraph.

Since new groups are formed by successful singletons, enough singletons must be in the population to form groups in the following round to replace the ones lost by invasion. However, the population cannot be comprised of entirely singletons; such a population would be unable to recover from a state of total defection without the formation of some cooperative groups to survive in a hostile environment. In a totally cooperating population, groups may be unneccesary; but such a population still runs the risk of falling into total defection where groups would prove beneficial.

We have argued that producing more singletons will lead to more cooperative groups (as defective singletons produce groups that die out, and cooperative singletons are copied more successfully). Combining the above we can now provide a reasonable explanation of the greater success of longer tags in promoting cooperation. The expression for M, however, tells us that similar effect can also be obtained by increasing mutation rate and population size. To validate our observations we ran experiments by varying one of N, μ, and l while keeping the other two constant. The base values used are $n = 100$, $\mu = 0.001$, $l = 32$. Each run consists of 100 generations.

Figure 6 shows the proportion of run during which percentage of cooperators in the population was above 90% with tag lengths used being powers of two between 2 and 64. It clearly denotes the beneficial effect of longer tags.

To observe the effect of higher mutation rates we plot the number of cooperators in the population in different generations of a run for mutation rates of 0.005 and 0.01 and $l = 8$, $N = 100$ in Figure 7. Each point is the average of 5 runs. We made a further change in that only the mutation rate on the tag bits was changed, the mutation rate of the strategy bit was kept at 0.001[1]. We find experimentally that increas-

[1] We believe that the mutation rate should be different on the tag bits compared to the strategy bit. In the current form the effective mutation rate on the strategy bit is $\frac{1}{l}$ times the mutation rate of the tag as the tag is l times longer. There is no reason for such an arbitrary dependence.

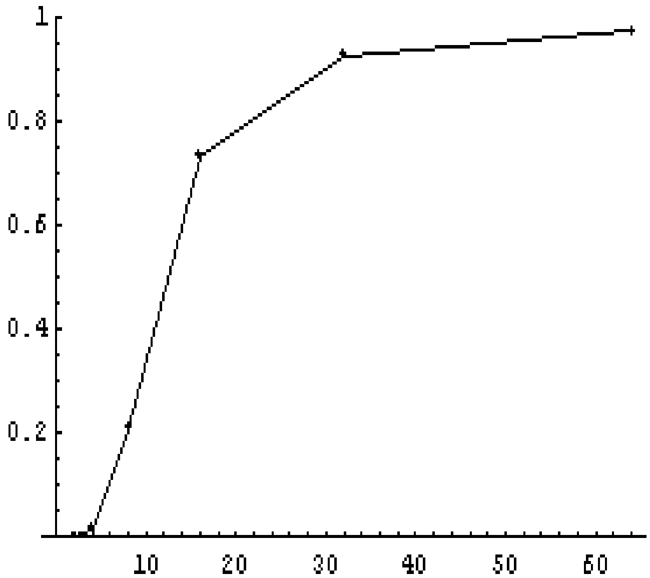

Figure 6: Proportion of run during which percentage of cooperators in the population was above 90% as a function of tag length.

ing mutation rate helps in improving the level of cooperation but not to the level of increasing tag length. Part of the discrepancy arises from the approximations used to derive the expected number of new singletons. Also, in light of the mechanisms for group change mentioned above, it is clear that muation rate affects 3 of the 4 mechanisms, while the other parameters have more limited effects. The trend, however, confirms our analysis that higher number of groups, in this case produced by higher mutation instead of higher tag length, helps in promoting and sustaining cooperation.

Discussion

In this paper, we have analyzed the population features influenced by varying tag lengths, which in turn gives rise to sustained cooperation in evolving population groups playing PD games. Our empirical observations are followed by some theoretical analysis which suggests that increasing mutation rate and population size can have similar effects. Such effects, though less dramatic than with increasing tag lengths, are observed in experiments.

There is a caveat to the above observations that should be highlighted here. We believe that ultimately if we vary only one of N, μ, and l, we will find a decrease in performance at high values. As N increases for a fixed l, mutations will just land an individual into another existing group and thus not help in creating new groups. As only μ increases, no groups, including cooperative groups, will be stable. As only l increases, interestingly, the same thing will happen as increasing μ as there is more chance a tag will get mutated. This suggests that ultimately we would need to use a mutation rate that is independent of the size of the tag space. We will run experiments to verify this conjecture.

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

References

Andreoni, J., and Varian, H. 1999. Preplay contracting in the prisoners' dilemma. *Proceedings National Academy of Sciences, USA* 96:10933–10938.

Axelrod, R. 1984. *The Evolution of Cooperation*. Basic Books.

Banerjee, B.; Sen, S.; and Peng, J. 2001. Fast concurrent reinforcement learners. In *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence*, 825–830.

Bowling, M., and Veloso, M. 2001. Rational and convergent learning in stochastic games. In *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence*, 1021–1026.

Brams, S. J. 1994. *Theory of Moves*. Cambridge University Press, Cambridge: UK.

Claus, C., and Boutilier, C. 1998. The dynamics of reinforcement learning in cooperative multiagent systems. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence*, 746–752. Menlo Park, CA: AAAI Press/MIT Press.

Dugatkin, L. A. 1997. The evolution of cooperation: four paths to the evolution and maintenance of cooperative behavior. *BioScience* 47(6):355–361.

Ferriere, R., and Michod, R. E. 1996. The evolution of cooperation in spatially heterogeneous population. *The American Naturalist* 147(5):692–718.

Hales, D., and Edmonds, B. 2003a. Can tags build working systems? from mabs to esoa. In *Engineering Self-Organising Systems*, Lecture Notes in AI-2977, 186–194. Springer Verlag.

Hales, D., and Edmonds, B. 2003b. Evolving social rationality for mas using "tags". In *Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems*, 497–503. Melbourne,Australia: ACM Press.

Holland, J. H.; Holyoak, K.; Nisbett, R.; and Thagard, P. 1986. *Induction: Processes of Inferences, Learning, and Discovery*. Cambridge, MA: MIT Press.

Holland, J. 1993. The effect of labels (tags) on social interactions. Technical Report Working Paper 93-10-064, Santa Fe Institute.

Holland, J. 2001. Exploring the evolution of complexity in signalling networks. *Complexity* 7(2):34–45.

Hu, J., and Wellman, M. P. 1998. Multiagent reinforcement learning: Theoretical framework and an algorithm. In Shavlik, J., ed., *Proceedings of the Fifteenth International Conference on Machine Learning*, 242–250. San Francisco, CA: Morgan Kaufmann.

Ito, A., and Yano, H. 1995. The emergence of cooperation in a society of autonomous agents – the Prisoner's Dilemma game under disclosure of contract histories. In *Proceedings of the First International Conference*

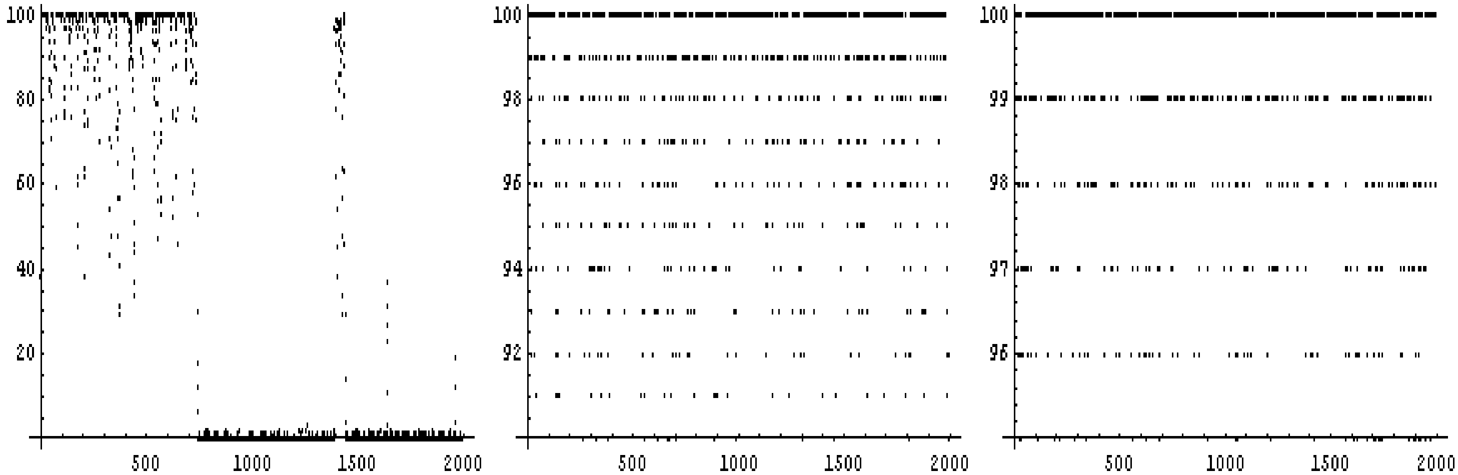

Figure 7: Number of cooperators over a run in experiments with mutation rate of 0.001 (left) 0.005 (center) and 0.01 (right).

on Multiagent Systems, 201–208. Menlo Park, CA: AAAI Press/MIT Press.

Littman, M. L., and Stone, P. 2001. Implicit negotiation in repeated games. In *Intelligent Agents VIII: AGENT THEORIES, ARCHITECTURE, AND LANGUAGES*, 393–404.

Littman, M. L. 1994. Markov games as a framework for multi-agent reinforcement learning. In *Proceedings of the Eleventh International Conference on Machine Learning*, 157–163. San Mateo, CA: Morgan Kaufmann.

Littman, M. L. 2001. Friend-or-foe q-learning in general-sum games. In *Proceedings of the Eighteenth International Conference on Machine Learning*, 322–328. San Francisco: CA: Morgan Kaufmann.

Myerson, R. B. 1991. *Game Theory: Analysis of Conflict*. Harvard University Press.

Nash, J. F. 1951. Non-cooperative games. *Annals of Mathematics* 54:286 – 295.

Posch, M. 1999. Win stay – lose shift strategies for repeated games: Memory length, aspiration levels, and noise. *Journal of Theoretical Biology* 198(2):183–195.

Riolo, R.; Cohen, M. D.; and Axelrod, R. 2001. Cooperation without reciprocity. *Nature* 414:441–443.

Riolo, R. 1997. 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*, 378–385. Morgan Kaufmann Publishers, Inc.

Sigmund, K., and Nowak, M. A. 1994. The alternating prisoner's dilemma. *Journal of Theoretical Biology* 38:262–275.

Stimpson, J. L.; Goodrich, M. A.; and Walters, L. C. 2001. Satisficing and learning cooperation in the prisoner's dilemma. In *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence*, 535–540.

Trivers, R. 1972. The evolution of reciprocal altruism. *Quarterly Review of Biology* 46:35–57.