Tag mechanisms evaluated for coordination in open multiagent systems

Isaac Chao[1], Oscar Ardaiz[2], Ramon Sanguesa[1]
[1] Informatics Department, Polytechnic University of Catalonia, Spain
{ichao, sanguesa}@lsi.upc.edu
[2] Informatics Department, Public University of Navarra, Spain
oscar.ardaiz@unavarra.es

Abstract. Tags are arbitrary social labels carried by agents. When agents interact preferentially with those sharing the same Tag, groups are formed around similar Tags. This property can be used to achieve desired group coordination by evolving agent's Tags through a group selection process. In this paper Tags performance is for the first time compared by simulation with alternative mechanisms for coordinated learning in multiagent systems populations. We target open systems, hence we do not make costly assumption on agent capabilities (rational or computational). It is a requirement that coordination strategies prove simple to implement and scalable. We build a simulator incorporating competition and cooperation scenarios modeled as one-shot repeated games between agents. Tags prove to be a very good coordination mechanism in both cooperation building for competitive scenarios and agent behavior coordination for fully cooperative scenarios.

Keywords: Tags, group selection, multiagent systems, coordination, prisoner's dilemma, cooperative games

1 Introduction

Tag-based coordination has already been evaluated for the Iterated Prisoner's Dilemma, (IPD) game and one-shot PD contexts. In this paper we contribute to evolve the sate of the art in Tag-based coordination in a twofold manner. First we build in a simulator the TagWorld [HALES00] model for the one-shot PD, and we evaluate the performance of Tags compared with alternative algorithms for multiagent systems (MAS) coordination. Second, we evaluate the Tag mechanism in a different scenario, a pure coordination game, accomplishing the same comparative study as for the PD. The results here can be useful for researchers aiming to use Tags as a coordination mechanism for engineering purposes.

An open system can be defined as one in which the structure of the system itself is capable of dynamically changing. The characteristics of such a system are that its components are not known in advance, can change over time and can consist of highly heterogeneous agents implemented by different people, at different times, with different software tools and techniques [SYCARA98]. Open systems incorporating high uncertainty are captured by the models used here, one-shot PD and one-shot pure coordination game.

1.1 Tag Mechanisms

The problem of coordination arises in (MAS) due to the distributed nature of the control exercised by the agents. Complexity and heterogeneity issues apply in open MAS, increasing enormously the costs of coordination. High dynamicity levels are also common in open systems, with agents entering and leaving in the system continuously (e.g. P2P systems), further complicating scalability. High levels of autonomy required by the agents are also in permanent conflict with coordination mechanisms. Distributed systems technologies (e.g. P2P, Grids) evolve more and more into open MAS systems, inheriting this problematic. We provide a coordination mechanism relevant for both selfish and cooperative agents in large, open MAS.

Recent trends in both distributed systems and MASs try to tackle the problem of coordination in MAS from a bottom-up point of view, studying the global properties that emerge from component/agent interaction. Tags can be used as powerful emergent coordination mechanism for MAS populations. Holland [HOLLAND93] first proposed the concept of Tags as markings or social cues attached to individuals (agents) and observable by others. Agents maintain and modify Tags on themselves and a team is formed by only collaborating with agents with the same Tag or some other condition. It is important not to mistake this concept of Tags with the collaborative Tagging phenomena [GOLDER05], so fashionable nowadays. The Tags we refer to in this paper are arbitrary social labels which do not convey any explicit meaning. Real-life examples of that kind of Tags are gang signals, native tongue and accent, skin color, etc. The Tags referred to in collaborative Tagging conveys meaning and are used to annotate semantically items to simplify further Tag-based search.

Riolo [RIOLO00] has described a number of Tagging approaches to address the IPD. Riolo outlines basic forms of Tagging: fixed-bias Tagging, variable-bias Tagging and evolved-bias Tagging. Tags promote the emergence of cooperation between agents even in the single round PD scenario [HALES00]. These techniques are attractive since they don't require centralized or third party reputation systems, the monitoring of neighbor behavior or the explicit programming of incentives. They also can be used in highly dynamic environments. The results from the single round PD scenario are especially interesting for the engineering of large-scale open distributed systems, since this situation is closer to a real highly dynamic open system, where heterogeneous agents are continuously entering and leaving the system, potentially accounting just for short term interactions.

In a basic Tag model simulation (TagWorld, [HALES00]), each agent maintains a strategy and a Tag (both can be initialized at random). Interaction involves pairs of agents playing a single round of PD. The Tag variable needs to have a quite big space available for variation (a full integer range suffices). This variable has no direct effect on the PD actions selected by the agent but is observable by all other agents. In this setting, a very simple algorithm is applied trough a number of rounds: First agents interact preferentially with other agents sharing the same Tag; then agents evolve following an evolutionary algorithm which preferentially reproduces agent's strategies that have collected bigger payoffs. Probabilistic mutation factors on both Tag and action variables are applied. The evolution of the population interacting following the prescribed mechanism precipitates a kind of "group selection" process in which those groups (each group being defined by a Tag) which are more cooperative tend to predominate but still die out as they are invaded by non cooperative agents. By constantly changing the Tag variable value (by reproduction of those with higher payoffs) the agents produce a dynamic process that leads to high levels of cooperative actions.

Notice that agents remain free to choose the actions cooperate or defeat. They still act in a selfish manner pursuing their own interest. However they commit to apply and respect the Tag algorithm. For more discussion elaborating on the implications of such requirement, see [MCDONALD05] and [ARTECONI07]. Extensive experimentation varying a number of parameters showed that for a big enough Tag space, high levels of cooperation quickly predominated in the population. Additionally, the fact that the system can recover from a state of total non-cooperative actions to almost total cooperative actions (under conditions of constant mutation) demonstrates high robustness. The Tag-based mechanism produces an efficient, scalable and robust solution based on very simple individual learning methods (modeled as reproduction and mutation).

1.2 Evaluating Tag Mechanisms

Provided the huge amount of work on MAS learning mechanisms, and the growing literature on Tag mechanisms, we state in this section motivation and targeted contributions of this paper. The paper aim is to further investigate the performance and the applicability of emergent coordination mechanisms based on Tags. Most of Tag-based related work is concerned with cooperation building in IPD settings (see all the papers in section 2). A few exceptions target applications of Tags in scenarios others than this. Notably, applications into realistic P2P scenarios are shown in the work by Hales [HALES04a], [HALES05], [HALES06]. Also other tentative applications have been query routing and processing for Peer-to-Peer web search [WEIKUM05], preventing free-riding in Grid Virtual Organizations coordination [CHAO06] and modeling the dynamics of firms [MOLLONA05].

We identify two important features missing in all this previous work. First, all the scenarios are targeting either cooperation building/free-riding controlling (all the IPD-based studies and the work on P2P systems by Hales) or competitive scenarios where agents have incentives to behave in opposition to the interest of the rest of the agents in the system (the rest of applications). What can we expect about the applicability of Tags into fully cooperative domains? Are Tags useful in those cases? Second, most of the Tag studies have been conduced relatively aside of the related body of work in MAS. Agent systems community has developed a number of learning algorithms which prove simple enough to be used by reactive agents in open systems environments. What knowledge about Tag mechanism can we derive by comparing its performance with existent MAS learning algorithms?

Provided the huge number of MAS learning mechanisms, it is useful to clarify the scope of the present work. The focus is extracting valuable information about the performance of Tag mechanism compared to very basic MAS learning algorithms. A good reason for choosing basic mechanisms is the fact that implementing Tags is extremely easy, and related computational cost very low, enabling for very large scalability. We purposely require the same feature from the rest of mechanisms selected for comparison. We explore the behavior of Tag mechanism in new domains, fully cooperative settings as opposed to those with agents having incentives to free-ride. Consequently, we concentrate on Tag mechanism evaluation in the simulation, and we do not cover in the same level of detail the rest of MAS learning mechanisms. We refer to related papers to cover in detail the variations on the rest of learning algorithms.

The results of the simulations show good performance for the Tag mechanism compared with alternatives, in both the cooperation and competition games. The contributions are twofold: We extend the knowledge of Tags in competitive settings (PD) by comparing its performance with other learning algorithms, and we identify novel scenario applications for the Tag mechanism, namely improved scalability of fully cooperative MAS settings. This opens a new path of applications of the Tag mechanisms, improving the scalability of cooperative learning agents, just by structuring the population of agents in groups and evolving these groups following the Tag algorithm (that is applying a process of group selection).

The rest of the paper structures as follows. Section 2 evaluates related work, relating its contributions to this paper. Section 3 details simulation environment, the Tag mechanism model and the alternative learning mechanisms simulated. Section 4 shows the core results: presents the experimental setup and the performance results on Tags compared with the alternative coordination mechanisms in both competitive and cooperative scenarios. A final subsection discussing the results and its applications is provided. Section 5 concludes the paper and outlines future work.

2 Related Work

In addition to the seminal research by Holland [HOLLAND93] and later by Riolo [RIOLO00] in the IPD setting and Hales [HALES00] in the one-shot PD setting, several recent papers have studied in-depth different aspects of Tag models. We present a summary of the most important conclusions and compare their contribution to ours.

In [HOWLEY05], the emergence of cooperation in simple Tag models incorporating IPD is studied by simulation. The results signal the importance of population viscosity (understood as static populations) in promoting cooperation between agents. Their simulation also proves that high Tag spaces are required for the emergence of cooperation. Although the number of total Tag values used by the agents fall significantly after the initial generations, the large number of Tags in the beginning is essential. The Tag mechanism has the ability to marginalize non cooperative behaviors over the initial populations

In the models from [MCDONALD05], it is confirmed by simulation that cooperation in Tag models is evolved based on fitness if sufficient number of new groups are created via mutation. 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. More interestingly, they build a partial theoretical characterization of this model, throwing the conclusion that Tag systems are merely promoting mimicry, rather than cooperation. In order to test this hypothesis, they build a model with agents playing a pure anti-coordination game instead of the PD. They conclude that in systems when cooperation requires complementary agents, Tags are not leading to cooperation. However, simulations by Hales [HALES04b] and Edmonds [EDMONDS06] show that some level of specialization can be derived between agents using Tags.

The two papers evaluated so far give a comprehensive evaluation on basic Tag mechanism behavior in PD settings (as pioneered by Riolo and Hales) but contrary to our paper, they do not provide any comparison with alternative learning mechanism. Also they focus on competitive scenarios and do not target fully cooperative games.

Research in [AVIV05] extends the use of Tags to interaction between groups, and not just to segment the population on groups of interacting agents as in previous models. The results of the simulations show that Tags incorporate some level of reciprocity between groups. In [ALKEMADE05] they present a Tag model incorporating sexual reproduction (recombination) of agents. Analyzing this model they find occasional formation of very stable cooperative societies, able to resists invasion of mimics (defecting agents with the Tag of a cooperative agent). Both models presents very interesting extensions of Tag mechanism, but no comparison with alternatives is present.

Perhaps the work which can be considered closer to this paper is presented in [COHEN99]. Here a wide comparison between many interaction-biasing processes is performed, which includes several topology-based, others based on random networks of neighbours and also Tags. The extensive simulation includes strategy variations and adaptation process variations for each of the interaction-biasing processes. The important result they achieve is that context-preservation, topologically-based or not, is essential in promoting cooperation. Tags are found to perform in-between full context preservation topological-based interaction processes and no context preservation process. In the Tag interaction process neighbours will tend to be chosen from a pool of like-Tagged agents, which is much smaller sized than the whole population of agents. This leads to an increasing probability of context preservation. This confirms the hypothesis by Howley [HOWLEY05] on viscosity requirements in Tag mechanisms, coming back to the findings on biological population's viscosity by Hamilton [HAMILTON71]. In contrast to our simulations, IPD is used here and not one-shot PD. Also, performance comparison is targeted in this paper specifically towards interaction processes analysis, rather than mechanisms understood as a system coordination artefact as in our case. This makes both comparative examinations complementary.

3 System model and learning mechanisms

3.1 System model: Cooperation and competition models

We use two fundamental games of game theory in order to represent the two basic scenarios: These are cooperation, where roughly individual and social welfare match, and competition of conflicting interests, where this is not necessarily the case (e.g. social dilemmas). The pure coordination game and the prisoner's dilemma respectively abstract those concepts.

The PD (Table 1) is a type of non-zero-sum game in which two players try to get rewards by cooperating with or betraying the other player. In the PD, cooperating is strictly dominated by defecting (i.e., betraying one's partner). Since in any situation playing defect is more beneficial than cooperating, all rational players will play defect (Nash Equilibrium). The unique Nash equilibrium for this game is a Pareto-suboptimal solution—that is, rational choice leads the two players to both play defect even though each player's individual reward would be greater if they both played cooperate. The challenge is to provide incentives in the repeated game for the agents to achieve the Pareto optimal solution maximizing population welfare, mutual cooperation. Let T stand for Temptation to defect, R for Reward for mutual cooperation, P for Punishment for mutual defection and S for Sucker's payoff. The following inequality must hold in a PD: $T > R > P > S$. If the game is iterated or repeated, the mutual cooperation total payment must exceed the temptation total payment: $2R > T + S$.

	P2 cooperates	P2 defects
P1 cooperates	R,R	S,T
P1 defects	T,S	P,P

Table 1: PD Game

The pure coordination game is symmetric, two player, two strategies, with payoff matrix as given in Table 2. In the coordination game the following holds: $A>C$ and $D>B$. Rational players will cooperate on either of the two strategies to receive a high payoff. Players in the game must agree on one of the two strategies in order to receive a high payoff. If the players do not agree, they receive a lower payoff. This game represents a common scenario in MAS systems when many agents' goals are to be aligned, leading to very suboptimal outcomes when this is not the case.

	P2 action 1	P2 action 2
P1 action 1	A,A	B,B
P1 action 2	C,C	D,D

Table 2: Coordination Game

3.2 Tag mechanism model

The proposed Tag model (see algorithm in figure 1) is close to TagWorld [HALES00], except a few details: We bootstrap the agents randomly into a number of groups (identified by a Tag) from the beginning, as opposed to Hales model where agents begin each one with a randomly given Tag. This initial bootstrapping is motivated to better model real scenarios where organizations are already in some specific configuration; as expected we did not find any impact for the two games analyzed of this initial bootstrapping. The second difference is that we explicitly forbid agent operation outside the group In the case of not finding a suitable mate in his own group the agent just skips operation in this round and goes directly to evolution phase. This should not make a big difference since isolated agents tend to migrate to other groups looking for bigger payoffs. This is motivated by the concern of keeping the group as the scope for agent operation. The last difference is that we mutate just Tags and do not introduce mutation in strategies. This is unnecessary for the emergence of cooperation/coordination (contrarily to Tag mutation). Equivalent effect to action mutation can be attained introducing noise in the game play, but this is not the point we target here. We refer to Hales [HALES00] for an analysis of strategy mutation impact and the robustness of the Tag mechanisms to it.

```
Bootstrap agents in groups
LOOP a number of rounds
    LOOP each group
        LOOP each agent in the group (operation phase)
        Interact with another agent from the group (i.e. same Tag)
            Collect Payoff
        ENDLOOP
    ENDLOOP
    LOOP each agent in the population (evolution phase)
            Select partner agents in the population
        If partner outperforms agent
            Then copy partner Tag (migrate to its group) and action
        Mutate: agent applies probabilistic Tag mutation
    ENDLOOP
ENDLOOP
```

Figure 1: Proposed Tag algorithm

The interaction involves Tag-biased mate selection (i.e. within the group) and bilateral playing of the game. The corresponding payoff is collected just by initiator agent. The evolution phase is ruled by evolutionary learning, replication of the fittest.

This is common in all Tag models from Riolo to Hales and others. Summarizing, the Tag mechanism model is very close to the one by Hales, except several details. The difference comes from the type of interactions; we will test the performance in pure coordination scenarios also, not only in social dilemma games such as PD.

Implicit assumptions of the Tag mechanism is that agents are capable of comparing utilities and relating them to actions, which is the case for the proposed games given that the agents know the payoff matrix. Different applications than the games tested here may require adequate services to assess/compare utilities. Another assumption is that a reliable discovery mechanism allows agents to locate other agents sharing the same Tag (i.e. belonging to the same group) and agents are able to communicate exchanging information. Scalability issues arising in practical applications from this assumption can be mitigated by sorting agents in groups following its Tags, reducing search spaces.

3.3. MAS learning algorithms selection

In order to achieve scalability in opens systems, we limit the type of coordination mechanism allowed to those meeting a set of conditions targeting scalability and practical issues: Being simple to implement (that means deploy and use) in real systems; not imposing computational or other expensive requirements on the agents; not requiring central coordination component (i.e. self-organizing). Properties are expected to emerge at the system level from individual agent's interaction. We exclude team learning and coalition formation literature from the set of mechanisms selected. Traditionally, most of the coalition formation algorithms are rather theoretical models based in deliberative agents. Limitations of these algorithms that render them inapplicable to large scale distributed systems are a high computational complexity, and unrealistic assumptions regarding the availability of information [SHEHORY04]. Some novel coalition formation mechanism could be incorporated in a simulation with more different scenarios [LERMAN00] [KRAUS03], though this is left for future work.

From the set of possible MAS mechanism to include in the simulation complying with the characteristics above, we have selected two important agent strategies emerging from game theory IPD tournaments, namely TFT and the Pavlov strategy (also called win stay, loose shift). We have introduced a simple Reinforcement Learning (RL) algorithm based on the extensions on the Pavlov strategy. Still in the camp of RL, we consider a basic Q-Learning algorithm. The last learning mechanism we have incorporated is a simple genetic algorithm for the whole population, with mimicry of the fittest alternative. It is important to notice that the proposed Tag algorithm incorporates the same evolutionary algorithm, with the only difference of splitting the population in groups. We use the following alternative learning mechanisms for the comparison:

- Generalized TFT, with random initialization of the action. This differs from original TFT definition [AXELROD81], but this is normal if we consider that the strategy is applied to the whole population, which would render a trivial game by starting all agents from cooperation. For the PD scenario TFT means player defect just if previous partner did so. Here TFT is applied to one-shot interactions with successive agents, which is a big difference from typical TFT applied in IPD tournaments. The purpose is to show how the population effectively converges to coordinated outcomes under this simple strategy, even in one-shot interactions. For the pure coordination game, TFT means to blindly copy previous partner agents, in the hope of having the whole population converging via this heuristic. Though this is of improbable success, we keep all learning mechanisms in both games to render some symmetry in the comparison.

- WSLS (Win stay, loose shift) or Pavlov strategy [NOWAK93]. Here an agent defects only if both the players did not agree on the previous move. It is a kind of TFT complement. The idea behind WSLS is to make agents responding partner agent moves following its experienced utilities, rather than partner's previous moves as in TFT. This makes the agents more reactive to their own experiences. The intuition is that emergence of coordination will be harder in one-shot scenarios than in IPD, since agents will have harder time relating utility to actions leading to payoff increases.

- RL: Basic Reinforcement Learning algorithm. Reinforcement Learning has been widely studied in agent theory (mostly in static environments) and in MAS settings. However the body of work in multi-agent RL is still small, contrasting with the literature on single-agent learning, as well as the literature on learning in game theory [SOHAM03]. It has been problematic extending convergence results to stochastic games. Consider for example the algorithm in [WAKANO01]: The RL formula for the agent internal state (Equation 1) can be explained as follows: We have for the agent an internal state, h, representing its satisfaction level. The learning player plays C (cooperation) when h>0, otherwise D (defection). If C is played and the resulting score f is larger than s (a fixed constant aspiration level), h increases, if smaller than s, then h decreases. Conversely successful D decreases h and vice-versa. The constant aspiration level s is taken in the interval $1 \leq s \leq 3$, in concrete we set s to 2. See more details about how this relates with the payoffs we use for the PD in [WAKANO01]. As for the pure coordination game, it is expected a good performance of this algorithm.

$$\Delta h = a \cdot \mathrm{sgn}(f - s) \cdot \mathrm{sgn}(h) \qquad (1)$$

- Q-Learning:

Q-learning [WATKINS89] is a form of RL algorithm that does not need a model of its environment and can be used on-line. Its convergence conditions have been widely studied, but convergence results in MAS have been always hard to achieve, rendering most of the advances to plausible heuristics [MUNOZ06]. We chose the simpler Q-learning available with typical values for the parameters including e-greedy selection. Again it is expected a good performance for the coordination game. For the PD, we have precedent studies showing good results for the IPD [SANDHOLM96]. However the performance in one-shot PD might differ from the one in IPD scenarios. We use the formula given in Equation 2 for Q-value updating. In this formula, α is the learning rate, the bigger the learning rate the more important is the impact of environment reinforcement. We set up a learning rate of 0.5. As for the Q-value selection, we use e-greedy policy with a probability of 0.1 of exploring randomly and a probability of 0.9 of exploiting the highest Q-value.

$$Q_{i,t+1} \leftarrow Q_{i,t} + \alpha(r - Q_{i,t}) \tag{2}$$

- Evolutionary (Evo): Just applies the most basic evolutionary learning, agents reproduce asexually copying the action played by the fittest partner they find. In our setting each agent chooses a partner randomly and copies its action if the agent outperforms him. In the same way that previous strategies (TFT and WSLS) and learning from reward mechanisms (basic RL, Q-learning) are used to decide and evolve the agent's next action, the evolutionary learning mechanisms (Evo and Tags) are used to derive next actions. What gets copied are agents actions (e.g. cooperate or defect for the PD, 0 or 1 for the pure coordination game) and not strategies (e.g. RL strategy, TFT strategy, etc). Since the evolutionary algorithm is also used in the Tag mechanism itself, the comparison can be used to further evaluate which building blocks from typical Tag mechanisms are promoting cooperation/coordination and how. For a full theoretical evaluation of this type evolutionary algorithm in a varied range of mutation levels, see [WILLENSDORFER05].

4 Experimental results and discussion

Given a population of N agents playing repeatedly one-shot PD and the pure coordination game, we want to compare the performance of several coordination mechanisms in order to maximize social welfare.

4.1 Experimental setup

We build a simulator in java enabling the implementation of different agent coordination protocols. Relevant parameters for the simulator are shown in table 3.
We fix a population of 100 agents, a representative size big enough to consider emergence in Tag-based coordination. Each experiment run is composed of 500 rounds to be able to average results over long runs. Tag mutation rate is fixed in the comparison with alternative MAS coordination mechanisms to m=0.01, a normal value in TagWorld, high enough to promote variability without fully destabilizing the systems. This is a typical value widely used in previous Tag mechanism studies.

NAME	VALUE
NUMBER OF ROUNDS	500 (fixed)
AGENT POPULATION	100 (fixed)
TAG MUTATION PROBAB m	0.001, 0.01, 0.1

Table 3: Experimental parameters

Agents are initialized with random actions in all cases, and given an initial Tag (in the case of the Tag mechanism). We instantiate the matrices for the PD and coordination games as shown in table 4. The payoffs for the PD are coherent with condition (2). The payoffs for the pure coordination game are the canonical.

PD	Cooperate	Defect	//	COORD	Action 1	Action 2
Cooperate	3,3	1,4	//	Strategy 1	1	-1
Defect	4,1	1,1	//	Strategy 2	-1	1

Table 4: Payoff Matrix instantiation

4.2 Performance comparison in the PD Game

For the PD game we measure two scores: Cooperation Level, the population percentage cooperating; varies on the range 0 to 1. Average Utility: Social Welfare; the population average utility. The bounds are 1(Sucker payoff) and 4(Temptation payoff). We are interested in the long term behavior of the Tag mechanism compared to the rest of mechanisms. We calculate on each round the average of the property (Cooperation Level or Average Utility) over the total number of rounds (500). We repeat the experiment for each parameters configuration 10 times and calculate the average with standard deviation as result of the experiment for each setting. The results are shown in table 5. We provide additional graphical display of a sample experiment run in figure 2.

PD GAME MECHANISM	COOP LEVEL MEAN	COOP LEVEL SD	/////	PD GAME MECHANISM	AV UTILITY MEAN	AV UTILITY SD
TFT	0,87	± 0,32	/////	TFT	2,72	± 0,42
WSLS	0,5	± 0,01	/////	WSLS	2,23	± 0,01
RL	0,20	± 0,23	/////	RL	1,72	± 0,48
QL	0,52	± 0,01	/////	QL	2,37	± 0,01
EVO	0,02	±0,01	/////	EVO	1,27	±0,03
TAGS (m=0.01)	0,98	± 0,01	/////	TAGS (m=0.01)	2,84	± 0,02

Table 5: Performance comparison for the PD.
Mean and standard deviation over 10 experiment runs

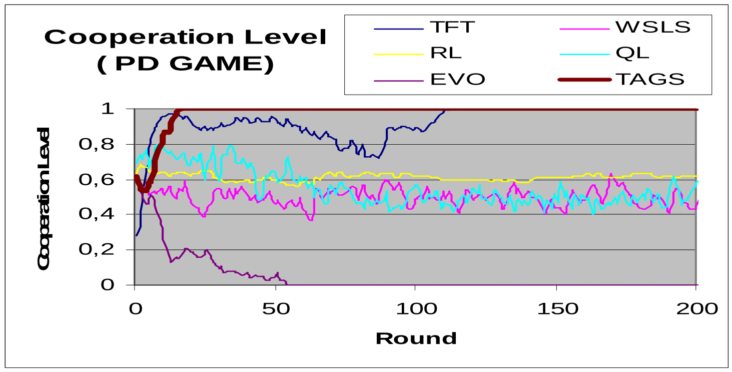

Figure 2: Performance results for a PD Game

From figure 2, we see in the long run similar performance of TFT and Tags in emerging total cooperation in the PD, but several rounds before in Tag-based models. The same stable pattern holds in rounds 200-500, hence we cut it. From the long run tabular results (Table5) we see a higher variance on the TFT mechanism in reaching cooperation, i.e. more unstable. Pavlov strategy (WSLS) is just able to maintain the initial random distribution of strategies, and never converges to any pure cooperation or defection. The simple RL mechanism performs badly; one-shot PD is harder scenario than IPD for coordination. Q-Learning mechanism maintains a stable equilibrium between cooperative and non cooperative agents but is not able to promote further cooperation. From the tabular results in table 5 we see an important difference on the variance, much higher in RL, proving a better stability on the Q-Learning. The evolutionary mechanism is not able to sustain cooperation. This confirms that niches formed by dividing the population in groups sharing Tag are essential in promoting cooperation (the context preservation referred in [COHEN99]).

4.3 Performance comparison in the Pure Coordination Game

For the Coordination Game we measure the Average Utility or Social Welfare. The bounds are 1(total coordination) and -1 (total uncoordinated behavior).

COORD GAME MECHANISM	AV UTILITY MEAN	AV UTILITY SD
TFT	-0,01	± 0,01
WSLS	-0,01	± 0,01
RL	-0,00	± 0,01
QL	0,05	± 0,01
EVO	0,94	±0,01
TAGS (m=0.01)	0,95	± 0,01

Table 6: Performance comparison for the Coordination Game.
Mean and standard deviation over 10 experiment runs

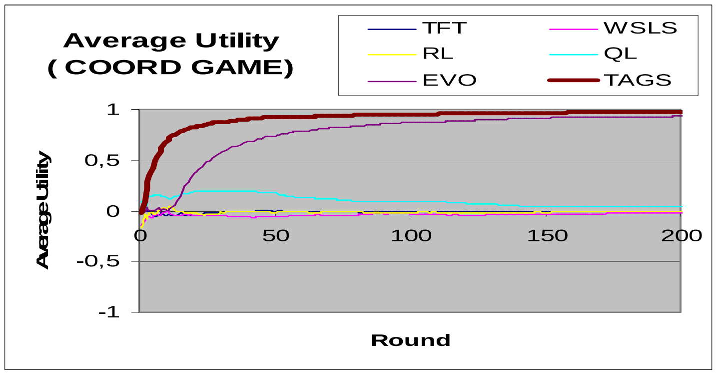

Figure 3: Performance results for the Coordination Game

Considering the pure coordination game (table 6, figure 3), TFT and WSLS are not helpful (as expected). The RL does not achieve any important improvement. The more elaborate Q-Learning algorithm is able to evolve a small level of coordination, This result is unexpected since reinforcement learning agents should achieve good performance in a so simple coordination game. However, again we recall multiagent RL literature on the controversy of convergence results in multiagent settings [BOWLING 00]. Our simulation shows two (basic) reinforcement learning mechanisms not converging on the PD and pure cooperation game scenarios. In contrast, Tags are able to coordinate the population. The evolutionary algorithm also performs very well. We conclude that it is the evolutionary aspect from Tag mechanisms which is mostly provoking the convergence of actions in fully cooperative domains.

4.4 Discussion and applications

We summarize here the most important conclusions of this comparative study. First, the Tag mechanism shows a very good performance in both scenarios, PD and pure coordination game. This confirms in broader MAS scope previous work by Riolo, Hales and others (see section 2) and extends the state of the art of Tag mechanism applied to fully cooperative domains. This expands Tag mechanisms from just social dilemmas and competitive settings to a full range of MAS applications, bridging its application to both competitive and cooperative multiagent learning fields [DURFEE94]. Second, we show how all the rest of simple MAS learning mechanism have worse performance in both competitive and cooperative scenarios, in the repeated one-shot games settings. It has been surprisingly found a bad performance of the reinforcement learning agents (simple RL and Q-Learning) in both games; especially surprising in the case of pure cooperation game. We propose two plausible reasons for this to happen: First the problems of RL with convergence in MAS settings anticipated in literature [BOWLING 00]. Second the fact that interactions are not iterated, rendering a much more unpredictable context for agents interactions. In such environment, simple model-free reinforcement learning agents get a hard time coordinating actions.

Tag mechanisms are a very promising mechanism for evolving swarms of agents into optimized outcomes. The applications of these results are varied in open MAS, in demanding scenarios where alternative learning mechanisms might fail or prove inapplicable due to computational requirements limitations. The most compelling are those envisaged for coordination purposes in Service Oriented Architectures [HUHNS05] and Next Generation Computational Grids [NGG06]. In this latter scenario, the management of first level system entities, Virtual Organizations (VOs), can be mapped directly to the concept of a group (identified by a Tag). In fact, most of the slight variations we introduced on the TagWorld model (see section 3.2) target the full alignment with the original VO concept from Grid computing domain. A research agenda has been outlined in [CHAO06], and partially addressed in the form of an application to automatic alignment of resource management policies in VOs [CHAO07a].

5 Conclusions and Future Work

In this paper Tags are for the first time compared by simulation with alternative mechanisms for coordinated learning in MAS populations. We target open MAS, hence we do not make costly assumptions on agent rational or computational capabilities. It is a requirement for us that coordination strategies prove simple and scalable. Tags are the simple, requiring for the agent just maintaining a marker visible to the rest of agents and show equal or better performance in the two games than any other of the mechanisms tested. Tags still remaining loosely coupled with the system, which enables, rather than substitution, for complementarities with other MAS coordination mechanisms. An example of this complementarily can be found in [CHAO07b], showing a mechanism which uses Tags optimizing decentralized Grid markets.

As we have reviewed in section 2, many related work has evaluated Tag mechanism internals for the IPD. Direct relation of the emergence of cooperation with mechanism parameters such as Tag space length and mutation rate is clearly identified and analyzed. However no prior study has evaluated how Tag mechanism performs in one-shot PD, compared to the alternative MAS learning algorithms. With the experiments from section 4, we have addressed this lack of comparability. We have found that Tags effectively work as indirect reciprocity enabler. Most of the algorithms had a hard time stabilizing their learning in the one-shot PD setting. Reciprocity built on repeated interaction with a same partner (as for TFT in IDP settings) or static contexts (as in single agent reinforcement learning) reveal crucial factors for the rest of the mechanism in order to achieve accurate learning. The basic strength of Tag mechanism is that it shows robust behavior in a demanding setting where other single MAS coordination mechanisms fail. Apart from reputation mechanisms, there are very few mechanisms addressing indirect reciprocity. Tags generate indirect reciprocity without the need of complex, system dependent reputation management, and perform well compared with alternatives. A partially theoretical analysis of all the mechanism enabling indirect reciprocity (including the group selection process behind Tag mechanism) can be found in [NOWAK06].

As for the pure coordination game, Tags achieve a good performance while the rest of the mechanisms again have problems coping with short-term interaction. The evolutionary mechanism is the exception, achieving a comparable performance. It is the evolutionary learning present in Tag mechanism the main driver of performance increases in this scenario. An important open issue in state-of-the-art Tag models is how to achieve coordination between complementary policies, which is difficult since Tags promote basically mimicry [MCDONALD05]. A solution to this issue may involve changes on the evolution of agent strategies when entering new groups, or alternatively trying more elaborate Tag similarity measures for interaction biasing.

Future work will comprise three basic areas. The first area is Tag mechanism extension to multiple Tags per agent and variations on the learning, using alternatives to evolutionary learning. Complementing the simulation approach with theoretical analysis, such as proposed in [GOTTS03] can elucidate better simulation scenarios, leading to improved accuracy. The second area consist on extending the comparison to more general coordination and organizational mechanisms, such as decentralized markets [EYMANN05] and tokens [XU06], and explore combinations between them (Tags are highly modular hence bear easy integration with existent mechanisms). Including scalable coalition-formation mechanisms [SHEHORY04] could also be done in this second comparison paper. A third area of improvement is on realistic models, beyond coordination games. In such models payoffs are based on real tasks execution modeling [GALSTYAN05] and not in abstract matrix. Deployment of Tags in a real prototype can lead to a performance evaluation of Tags in real networks settings.

References

[ALKEMADE05] Floortje Alkemade, D. D. B. van Bragt, Johannes A. La Poutré: Stabilization of Tag-mediated interaction by sexual reproduction in an evolutionary agent system. Inf. Sci. 170(1): 101-119 (2005)

[ARTECONI07] Arteconi, S.; Hales, D.; Babaoglu, O. (2007) Greedy Cheating Liars and the Fools Who Believe Them. In Engineering Self-Organising Systems, Proceedings of the 4th International Workshop, ESOA 2006, Hakodate, Japan, May 9, 2006, Revised and Invited Papers. Brueckner, S.; Hassas, S.; Jelasity, M.; Yamins, D. (Eds.), Lecture Notes in Computer Science , Vol. 4335, Springer

[AVIV05] Using Tags to Evolve Trust and Cooperation Between Groups, Aviv Zohar and Jeffrey S. Rosenschein. The Fourth International Joint Conference on Autonomous Agents and Multiagent Systems, July 2005, Utrecht, The Netherlands, pp. 1199-1200

[AXELROD81] Axelrod, Robert .The Evolution of Cooperation. Science, 211(4489):1390-6 (1981).

[BOWLING 00] Michael Bowling. Convergence problems of general-sum multia-gent reinforcement learning. In Proceedings of the Seventeenth International, Conference on Machine Learning, pages 89–94, 2000.

[CHAO06] Isaac Chao, Oscar Ardaiz, Ramon Sanguesa . Tag Mechanisms Applied to Open Grid Virtual Organizations Management, in Proceedings of the Joint Smart Grid Technologies (SGT) and Engineering Emergence for Autonomic Systems (EEAS)

Workshop, Editors: R. Anthony, A. Butler, M. Ibrahim, T. Eymann, and D. J. Veit, Dublin, Ireland, pp 22-29, 2006
[CHAO07a] Isaac Chao, Oscar Ardaiz, Ramon Sanguesa. A Group Selection Pattern for agent-based Virtual Organizations coordination in Grids, accepted to the International Conference on Grid computing, High-Performance and Distributed Applications (GADA'07), Vilamoura, Algarve, Portugal, Nov29 - 30, 2007.
[CHAO07b] Isaac Chao, Oscar Ardaiz, Ramon Sanguesa. A Group Selection Pattern Optimizing Job Scheduling in Decentralized Grid Markets. Poster accepted to the International Conference on Grid computing, High-Performance and Distributed Applications (GADA'07), Vilamoura, Algarve, Portugal, Nov29 - 30, 2007.
[COHEN99]Cohen, M., R. Riolo, and R. Axelrod(1999): "The emergence of social organization in the Prisoner's Dilemma: how context-preservation and other factors promote cooperation," Santa Fe Institute Working Paper 99-01-002
[DURFEE94] E.H. Durfee and J.S. Rosenschein, “Distributed Problem Solving and Multi-Agent Systems: Comparisons and Examples,” Proc. 13th Int’l Distributed Artificial Intelligence Workshop, pp. 94– 104, 1994
[EDMONDS06] The Emergence of Symbiotic Groups Resulting from Skill-Differentiation and Tags, Bruce Edmonds, JASSS 9(1), 2006/1/31
[EYMANN05] Torsten Eymann, Michael Reinicke, Werner Streitberger, Omer Rana, Liviu Joita, Dirk Neumann, Björn Schnizler, Daniel Veit, Oscar Ardaiz, Pablo Chacin, Isaac Chao, Felix FreiTag, Leandro Navarro, Michele Catalano, Mauro Gallegati, Gianfranco Giulioni, Ruben Carvajal Schiaffino, Floriano Zini, “Catallaxy-based Grid Markets”, Multiagent and Grid Systems Issue: Volume 1, Number 4 / 2005 Pages: pp. 297 - 307 Special Issue: Smart Grid Technologies & Market Models,Guest Editor(s): Daniel J. Veit, Torsten Eymann, Nicholas R. Jennings, Jörg P. Müller
[GALSTYAN05] Aram Galstyan, Karl Czajkowski and Kristina Lerman, “Resource Allocation in the Grid with Learning Agents,” Journal of Grid Computing 3(1–2):91 –100, 2005
[GOLDER05] Golder, Scott A. Huberman, Bernardo A. (2005). "The Structure of Collaborative Tagging Systems." Information Dynamics Lab, HP Labs. Visited November 24, 2005.
[GOTTS03] Agent-based simulation in the study of social dilemmas., Gotts, N.M., Polhill, J.G. and Law, A.N.R., (2003) Artificial Intelligence Review, 19, 3-92
[HALES00] Hales, D. (2000) Cooperation without Space or Memory: Tags, Groups and the Prisoner's Dilemma. In Moss, S., Davidsson, P. (Eds.) Multi-Agent-Based
[HALES04a] Hales, D. (2004) From Selfish Nodes to Cooperative Networks – Emergent Link-based Incentives in Peer-to-Peer Networks. In proceedings of The Fourth IEEE International Conference on Peer-to-Peer Computing (p2p2004), 25-27 August 2004, Zurich, Switzerland. IEEE Computer Society Press
[HALES04b] Hales, D. (2004) The Evolution of Specialization in Groups. In Lindemann, G., Moldt, D. and Paolucci, M., (Eds.) Regulated Agent-Based Social Systems - 1st Intnerational Workshop, RASTA 2002 Bologna, Italy, July 2002. Revised Selected and Invited Papers. Lecture Notes in Artificial Intelligence 2934. Berlin: Springer-Verlag.
[HALES05] Hales, D. & Patarin, S. (2005) Feature: Computational Sociology for Systems "In the Wild": The Case of BitTorrent. IEEE Distributed Systems Online, vol. 6, no. 7, 2005
[HALES06] Hales, D. and Babaoglu, O. (2006) Towards Automatic Social Bootstrapping of Peer-to-Peer Protocols. In ACM SIGOPS Operating Systems Review (Special Issue on Self-Organizing Systems), vol. 40, no. 3, July 2006.
[HAMILTON71] W.D. Hamilton, Man and Beast: Comparitive Social Behaviour, Smithsonian Press, Washington DC. eds., J.F. Eisenberg, W.S. Dillon, 1971.
[HOLLAND93] J. Holland. The effects of labels (Tags) on social interactions. Working Paper Santa Fe Institute 93-10-064, 1993.

[HOWLEY05] E. Howley and C. O'Riordan, The Emergence of Cooperation among Agents using Simple Fixed Bias Tagging, IEEE Congress on Evolutionary Computation, September 2-5, 2005

[HUHNS05] Research Directions for Service-Oriented Multiagent Systems. M.N. Huhns, M.P. Singh, M. Burstein, K. Decker, E. Durfee, T. Finin, L. Gasser, H. Goradia, N. Jennings, K. Lakkaraju, H. Nakashima, H.V.D. Parunak, J.S. Rosenschein, A. Ruvinsky, G. Sukthankar, S. Swarup, K. Sycara, M. Tambe, T. Wagner, and L. Zavala

[KRAUS03] Kraus, S., Shehory, O., Taase, G.: Coalition formation with uncertain heterogeneous information, Proc. Of AAMAS-03, pages 1–8, Melbourne, Australia, 2003.

[LERMAN00] Lerman, K., Shehory, O.: Coalition formation for large-scale electronic markets, Proc. of ICMAS-00, pages 167–174, Boston, MA, 2000.

[MCDONALD05] Austin McDonald and Sandip Sen, "The Success and Failure of Tag Mediated Evolution of Cooperation" Woking Notes of the AAMAS-05 Workshop on Learning and Adaptation in MAS (LAMAS) 155-164

[MOLLONA05] Mollona, E. & Hales, D. (in press) Modeling Firm Skill-Set Dynamics as a Complex System. Proceedings of the European Conference on Complex Systems (ECCS'05), Nov. 14th, Paris, France. Publishers: i6doc, Belgium

[MUNOZ06] Jose Enrique Munoz de Cote, Alessandro Lazaric, Marcello Restelli: Learning to cooperate in multi-agent social dilemmas. AAMAS 2006: 783-785

[NGG06] Next Generation Grids Expert Group Report 3, Future for European Grids: GRIDs and Service Oriented Knowledge Utilities.

[NOWAK06] Nowak MA (2006). Five rules for the evolution of cooperation. Science 314: 1560-1563

[NOWAK93] Nowak, M., and Sigmund, K. 1993. A strategy of winstay,lose-shift that outperforms tit-for-tat in the prisoner'sdilemma game. Nature 364:56–58.

[RIOLO00] R. Riolo. The efects of Tag-mediated selection of partners in evolving populations playing the iterated prisoners dilemma. Nature 414, pages 441–443, 2000.

[SHEHORY04] Coalition Formation: Towards Feasible Solutions. Fundamenta Informaticae, Vol. 63, No. 2 - 3. (January 2004), pp. 107-124

[SOHAM03] Shoham, Y., Grenager, T., Powers, R.: Multi-agent reinforcement learning: A critical survey. Tech.rep., Stanford University (2003)

[SYCARA98] Multiagent Systems K. Sycara, AI Magazine, Vol. 10, No. 2, 1998, pp. 79-93

[WAKANO01] Wakano, J. Y. and Yamamura, N. (2001) A Simple Learning Strategy that Realizes Robust Cooperation Better than Pavlov in Iterated Prisoners' Dilemma. J. Ethology 19: 9-15

[WATKINS89] Watkins,C,1989.Learning from Delayed Rewards,Thesis,University of Cambidge,England

[WEIKUM05] Gerhard Weikum, Peter Triantafillou, David Hales, Christian Schindelhauer (in press) Towards Self-Organizing Query Routing and Processing for Peer-to-Peer WebSearch. Proceedings of the European Conference on Complex Systems (ECCS'05), Nov. 14th, Paris, France. Publishers: i6doc, Belgium

[WILLENSDORFER05]Willensdorfer M, MA Nowak (2005). Mutation in evolutionary games can increase average fitness at equilibrium. J theor Biol 237: 355-362