Emergence of Norms for Social Efficiency in Partially Iterative Non-Coordinated Games

(Extended Abstract)

Toshiharu Sugawara
Department of Computer Science and Engineering
Graduate School of Waseda University
3-4-1 Okubo, Shinjuku
Tokyo 169-8555, Japan
sugawara@waseda.jp

ABSTRACT

We discuss the emergence of social norms for efficiently resolving conflict situations through reinforcement learning and investigate the features of the emergent norms, where conflict situations can be expressed by non-cooperative payoff matrix and will remain if they fail to resolve the conflicts.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Multiagent systems*

General Terms

Experimentation

Keywords

Norm, Conflict, Coordination, Reinforcement Learning

1. INTRODUCTION

Facilitation of coordination and conflict resolution is an important issue in multi-agent systems. One method to cope with these problems is to use social laws or social norms that all agents are expected to follow. In this paper, we discuss whether the conventions which are the special form of norms can emerge by reinforcement learning and based on the payoff matrices that characterize the participating agents.

There are a number of studies on learning-based norm emergence such as in [2, 3], where agents individually learn, through interactions with others, the identical conventions that maximize their payoffs. For example, [2], uses coordination games that have simple but multiple equilibria, whereby all agents evolve a policy to select one of the equilibria.

On the other hand, our research concerns competitive or conflicting situations that can be expressed as a two-player game. In addition, the game is iterated if the agents fail to resolve the situation. Thus, agents want to make the society more efficient by using norms.

Cite as: Emergence of Norms for Social Efficiency in Partially Iterative Non-Coordinated Games (Extended Abstract), Toshiharu Sugawara, *Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, Tumer, Yolum, Sonenberg and Stone (eds.), May, 2–6, 2011, Taipei, Taiwan, pp. 1193-1194.

The aim of this paper is to investigate the question of whether norms that lead efficient conflict resolutions in the society emerge as a result of reinforcement learning. Although agents act and learn according to their own payoff matrices, they may have different matrices in each experiment; thus, the average payoffs that all the agents gain can't be compared with each other. Some agents cannot take the obvious best action that may lead to zero or negative payoffs because of conflicts. However, by taking a less than best action, they may be able to create an efficient society and, as the result, yield better payoffs in the end. In such situations, we want to investigate how changing the agent type affects norm emergence and the efficiency of the resulting societies. Our results indicate that agents having explicit orders of actions can evolve stable social norms, whereas those that are not willing to give the other an advantage (negative payoffs) cannot evolve stable norms.

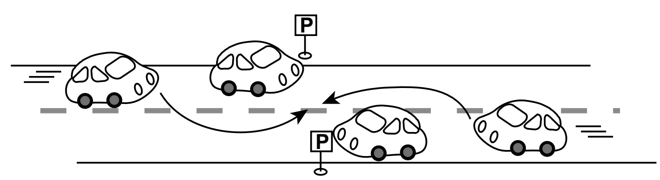

Figure 1: Narrow Road Game.

2. MODEL AND PROBLEM

2.1 Narrow Road Game in Agent Society

We consider a modified version of the *narrow road game* (MNR game) [1] in which car agents encounter the situation shown in Fig. 1. This is a two-player game, more precisely a sort of Markov game, expressed by the following payoff matrix where the agents take one of two actions, p (proceed) or s (stay):

$$\begin{array}{c} \\ p \\ s \end{array} \begin{array}{c} \quad p \quad\;\; s \leftarrow \text{Actions of the adversary agent.} \\ \begin{pmatrix} -5 & 3 \\ -0.5 & 0 \end{pmatrix} \end{array} \qquad \text{(M1)}$$

The agents having matrix (M1) receives -5 (maximum penalty) if their action is (p, p) because they have to go back to escape the deadlock. However, (s, s) does not induce any benefit or penalty because no progress occurs (later, we introduce a small penalty for (s, s)).

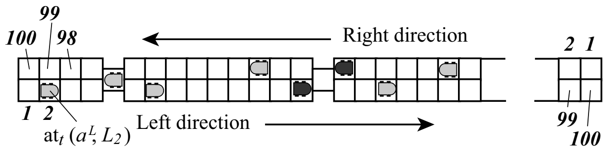

Figure 2: Narrow Road Game.

We consider that agents in two parties A_L and A_R, which are the disjoint sets of agents, play the MNR game. We also assume that $A = A_L \cup A_R$ is the society of agents. The two-lane road, as shown in Fig. 2, is one in which agents in A_L (A_R) move forward in the left (right) lane. The road has a number of narrow parts where left and right lanes converge into one. In this environment, two agents $a_i^L \in A_L$ and $a_j^R \in A_R$ play the narrow-road game when they are on ether side of a narrow part. We assume that this road has a ring structure.

2.2 Emergence of Norms and Payoff Matrices

We investigated how agents learn the norms for the MNR games by reinforcement learning and how their society becomes more efficient as a result of the emergent norms. We expect that the consistent joint norm in agents in A_L (or A_R) emerges.

To introduce some kinds of agents in this game, we define addition four payoff matrices that characterize the agents:

($M2$) Moderate $\begin{pmatrix} -5 & 3 \\ 0.5 & 0 \end{pmatrix}$ ($M3$) Selfish $\begin{pmatrix} -5 & 3 \\ -0.5 & -0.5 \end{pmatrix}$

($M4$) Generous $\begin{pmatrix} -5 & 3 \\ 3 & -0.5 \end{pmatrix}$ ($M5$) Self-centered $\begin{pmatrix} -5 & 3 \\ -5 & -0.5 \end{pmatrix}$

Note that we call an agent characterized by matrix M1 *normal*. An agent has only one payoff matrix.

Matrix M2 characterizes a *moderate* agent whose payoff of (s, p) is 0.5 (positive); it may be able to proceed the next time. The *selfish* (or *self-interested*) agent is characterized by M3 which has a positive payoff only when it can proceed the next time. (Joint action (s, s) also induces a small penalty because it is the waste of time). The *generous* agent defined by M4 does not mind if its adversary proceeds first (it can proceed the next time if the game is over). This matrix defines the coordination game and has two obvious equilibria [2] if this is not a Markov game. The *self-centered* agent characterize by (M5) is only satisfied when it can proceed and is very unhappy if the adversary goes ahead. Matrix M5 has the obvious best action p if the game is not iterative.

3. EXPERIMENT – IMPROVEMENT OF SOCIAL EFFICIENCY

We assume that the populations of both parties $|A_L|$ and $|A_R|$ are 20, the road length l is 100 and there are four narrow parts along the road (the positions are random). All agents in A_L (A_R) are randomly placed on the left (right) lane except the narrow parts. The data shown in this paper are the average values of 1000 trials.

We examine a number of cases, but here we will show the result when the society consists of homogeneous agents. We compared the average *go-round times* (AGRT) of the societies, where go-round time is the time required to come back to the start position. The results are shown in Fig. 3.

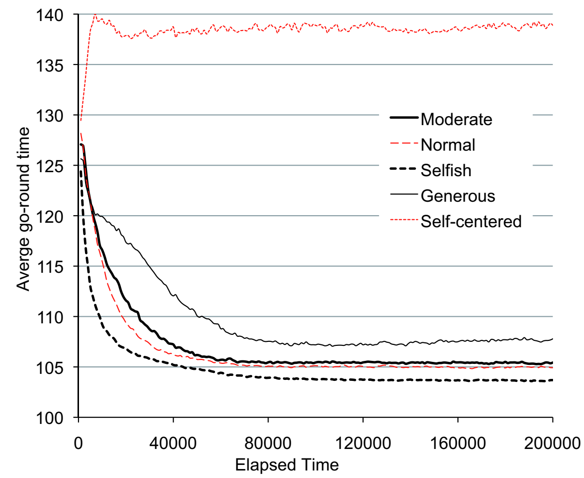

Figure 3: Average go-round time (AGRT).

This figure indicates that the AGRT values become smaller in all societies except the self-centered one. Because a smaller AGRT means that conflicts can be resolved more quickly, we can say that the society become more efficient by reinforcement learning.

4. CONCLUDING REMARK

We are interested in the emergence of social norms (conventions) that may incur a certain cost/penalty to a number of agents but are beneficial to the society as a whole. This kind of norm plays a significant role in conflicting situations.

Our results showed that selfish agents, which have a large positive payoff for its own advantage and a small negative payoff for other's advantage, can evolve stable social norms. However, they cannot evolve norms if they also have a large negative payoff for the adversary's advantage.

5. REFERENCES

[1] K. Moriyama and M. Numao. Self-Evaluated Learning Agent in Multiple State Games. In *Proceedings of the 14th European Conference on Machine Learning, ECML-2003 (LNCS 2837)*, pages 289–300. Springer, 2003.

[2] P. Mukherjee, S. Sen, and S. Airiau. Norm emergence under constrained interactions in diverse societies. In *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pages 779–786. IFAAMAS, 2008.

[3] S. Sen and S. Airiau. Emergence of Norms Through Social Learning. In *International Joint Conference on Artificial Intelligence (IJCAI-07)*, pages 1507–1512, 2007.