IGI PUBLISHING
701 E. Chocolate Avenue, Suite 200, Hershey PA 17033-1240, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.igi-global.com

ITJ 4731

This paper appears in the publication, International Journal of Agent Technologies and Systems, Volume 1, Issue 2
edited by **Goran Trajkovski © 2009, IGI Global**

Norm Emergence with Biased Agents

Partha Mukherjee, University of Tulsa, USA

Sandip Sen, University of Tulsa, USA

Stéphane Airiau, University of Amsterdam, The Netherlands

ABSTRACT

Effective norms can significantly enhance performance of individual agents and agent societies. We consider individual agents that repeatedly interact over instances of a given scenario. Each interaction is framed as a stage game where multiple action combinations yield the same optimal payoff. An agent learns to play the game over repeated interactions with multiple, unknown, agents. The key research question is to find out whether a consistent norm emerges when all agents are learning at the same time. In real-life, agents may have pre-formed biases or preferences which may hinder or even preclude norm emergence. We study the success and speed of norm emergence when different subsets of the population have different initial biases. In particular we characterize the relative speed of norm emergence under varying biases and the success of majority/minority groups in enforcing their biases on the rest of the population given different bias strengths. [Article copies are available for purchase from InfoSci-on-Demand.com]

Keywords: *Autonomous Agents; Emergence; Simulation; Time; Time-Ecologies*

INTRODUCTION

Recent literature in multiagent systems show a significant increase in interest and research on normative systems which are defined as (Boella, Torre, & Verhagen, 2008):

A normative multiagent system is a multiagent system organized by means of mechanisms to represent, communicate, distribute, detect, create, modify, and enforce norms, and mechanisms to deliberate about norms and detect norm violation and fulfillment.

Norms or conventions routinely guide the choice of behaviors in human societies and plays a pivotal role in determining social order (Hume, 1978). Conformity to norms reduces social frictions, relieves cognitive load on humans, and facilitates coordination. This is because individuals conform to norms assuming others are going to do the same (Lewis, 1969). Typically norms facilitate social interactions by enabling interacting agents choose actions without social enforcement, that result in coordinated behavior, e.g., driving on appropriate sides of the road. Norms may be adhered to in human societies because they facilitate the functioning of individuals, or because of the threat of social disapproval (Posner, 2000) or acceptance by individuals of desired conduct (Elster, 1989).

Computational agents, too, often have to coordinate their actions and adoption and adherence to norms can improve the efficiency of agent societies. A large class of interactions between self-interested agents (players) can be formulated as stage games with simultaneous moves made by the players (Genesereth, Ginsberg, & Rosenschein, 1986). Such stage games often have multiple equilibria (Myerson, 1991), which makes coordination uncertain. While *focal points* (Schelling, 1960) can be used to disambiguate such choices, they may not be available in all situations. Norms can also be thought of as focal points evolved through learning (P. H. Young, 1996) that reduce disagreement and promote coherent behavior in societies with minimal oversight or centralized control (Coleman, 1987). Norms can therefore have economic value to agents and help improve their efficiency. Norms in human societies, however, can also prevent flexibility and do not necessarily efficiency. In addition, different societies may evolve different norms to solve the same coordination problem, norms may shift over time, and can have varying degrees of stability (H. P. Young, 2008). Hence, the systematic study and development of robust mechanisms that facilitate emergence of stable, efficient norms via learning in agent societies promises to be a productive research area that can improve coordination in and thereby functioning of agent societies.

Establishment of social norms may come about by top-down influences like official edicts and role models, bottom-up processes driven by local customs, and lateral diffusion of established norms between related interaction types (H. P. Young, 2008). Most research on norms in multiagent systems focus on the *legalistic view* where norms are used to shape the behavior of open systems without using sanctions to enforce desirable behavior. In this approach norms are typically logically specified using a normative language (García-Camino, Rodríguez-Aguilar, & Sierra, 2005) from which rules of behavior can be automatically derived (Silva, 2008). Our approach to norm emergence from personal interactions is based on the *interactionist view* which adopts a bottom-up view of individual adoption of norms because of alignment of goals and utilities between agents in a population (Castelfranchi, 1998, 2003). In addition to the appeal of distributed, rather than centralized approach, the process of norm emergence can also facilitate efficiency and promote fairness (Binmore, 1994, 2005).

While researchers have studied the emergence of norms in agent populations, they typically assume access to significant amount of global knowledge (Epstein, 2001, Kandori & Rob, 1995, P. H. Young, 1993, 1996). For example, all of these models assume that individual agents can observe sizable fraction of interactions between other agents in the environment. While these results do provide key insights into the emergence of norms in societies where the assumption of observability holds, it is unclear if and how norms will emerge if all interactions were private, i.e., not observable to any other agent not involved in the interaction.

To study the important phenomenon of emergence of social norms via private interactions, we use the following interaction framework. We consider a population of agents, where, in each interaction, each agent is paired with another agent randomly selected from the population. Each agent then is learning concurrently over repeated interactions with randomly selected members from the population. We refer to this kind of learning *social learning* to distinguish from learning in iterated games against the same opponent (Fudenberg & Levine, 1998). Our experiments involve symmetrical games with multiple pure-strategy equilibria with the same payoff.

In previous work on learning in games, the opponent is fixed. In our *social learning framework* (Mukherjee, Sen, & Airiau, 2007, Sen & Airiau, 2007), the opponent is unknown and different at each iteration. In past work (Mukherjee et al.., 2007, Sen &

Related Content