A categorization of simulation works on norms

Bastin Tony Roy Savarimuthu and Stephen Cranefield

Department of Information Science, University of Otago, Dunedin, P O Box 56, Dunedin, New Zealand (tonyr,scranefield)@infoscience.otago.ac.nz

Abstract. In multi-agent systems, software agents are modelled to possess characteristics and behaviour borrowed from human societies. Norms are expectations of behaviours of the agents in a society. Norms can be established in a society in different ways. In human societies, there are several types of norms such as moral norms, social norms and legal norms (laws). In artificial agent societies, the designers can impose these norms on the agents. Being autonomous, agents might not always follow the norms. Monitoring and controlling mechanisms should be in place to enforce norms. As the agents are autonomous, they themselves can evolve new norms while adapting to changing needs. In order to design and develop robust artificial agent societies, it is important to understand different approaches proposed by researchers by which norms can spread and emerge within agent societies. This paper makes two contributions to the study of norms. Firstly, based on the simulation works on norms, we propose a life-cycle model for norms. Secondly, we discuss different mechanisms used by researchers to study norm creation, spreading, enforcement and emergence.

1 Introduction

In human societies, norms have played an important role in governing behaviour of the individuals in a society. Norms are the societal rules that govern the prescription and proscription of certain behaviour . Norms improve cooperation [1] and coordination among agents [2]. Norms reduce the amount of computation required by the agents [3] as the agents do not have to search their entire state space if they were to follow norms.

Artificial agent societies are societies in a networked environment where various agents share a virtual space and perform certain actions in a particular context (e.g. auctions). These agent societies are modelled using some of the social constructs borrowed from the human society. There have been two approaches for building normative behaviour in an agent. The first approach is the prescriptive approach where an institutional mechanism specifies how the agents should behave. The second approach is the bottom-up approach that could be used in open environments by employing mechanisms that can help a norm to emerge and govern the behaviour of an agent.

The advent of digital virtual environments such as Second Life [4] call for a distributed approach to norm spreading and emergence. Centralized policing

Dagstuhl Seminar Proceedings 09121
Normative Multi-Agent Systems
http://drops.dagstuhl.de/opus/volltexte/2009/1905

mechanism for such digital societies would be expensive from the view point of computation required due to the explosion of the states of the agents. It is impossible to monitor and control millions of agents assuming numerous roles through a centralized enforcer. A distributed approach to norms addresses these problems. Both of these approaches have been addressed by researchers and the current works focus on the issues associated with the distributed approach.

A "norm capable" agent society is the one that is able to generate, distribute, enforce and modify norms. Building robust agent societies that can create and evolve norms is important because the framework that helps in recognizing these norms will also be helpful for the agents to dynamically change these norms if situations warrant it. A good approach to test models of norm capable societies are simulations. So, a first step towards building such norm capable societies is to understand the existing simulation works on norms. To that extent, based on the simulation works on norms, we propose a life-cycle model for norms in the first part of the paper and in the second part of the paper we categorize the research works on norms based on the mechanisms employed by each of works.

2 What are norms?

Norms are expectations of an agent about the behaviour of other agents in the society. The human society follows norms such as tipping in restaurants and exchange of gifts at Christmas. Norms have been so much a part of different cultures, it is not surprising that it is an active area of research in a variety of fields including Sociology, Economics, Biology and Computer Science. Social norms have been of interest to multi-agent researchers since the early nineties. Norms are of interest to multi-agent system (MAS) researchers as they help in sustaining social order and increase the predictability of behaviour in the society. However, software agents tend to deviate from these norms due to their autonomy. So, the study of norms has become crucial to MAS researchers as they can build robust multi-agent systems using the concept of norms and also experiment with how norms evolve and adapt in response to environmental changes.

Due to multi-disciplinary interest in norms, several definitions for norms exist. Habermas [5], a renowned sociologist, identified norm-regulated actions as one of the four action patterns in human behaviour. A norm to him means *fulfilling a generalized expectation of behaviour*, which is a widely accepted definition for social norms. Ullman-Margalit [6] describes a social norm as a prescribed guide for conduct or action which is generally complied with by the members of the society. She states that norms are the resultant of complex patterns of behaviour of a large number of people over a protracted period of time. Coleman [7] describes *"I will say that a norm concerning a specific action exists when the socially defined right to control the action is held not by the actor but by others"*. Elster notes the following about social norms [1]. *"For norms to be social, they must be shared by other people and partly sustained by their approval and disapproval. They are sustained by the feelings of embarrassment,anxiety, guilt and shame that a person suffers at the prospect of violating them. A person obeying*

a norm may also be propelled by positive emotions like anger and indignation ... social norms have a grip on the mind that is due to the strong emotions they can trigger".

Researchers have divided norms into different categories. Tuomela [8] has grouped norms into the following categories.

- r-norms (rule norms)
- s-norms (social norms)
- m-norms (moral norms)
- p-norms (prudential norms)

Rule norms are imposed by an authority based on an agreement between the members (e.g. one has to pay taxes). Social norms apply to large groups such as a whole society (e.g. one should not litter). Moral norms appeal to one's conscience (e.g. one should not steal or accept bribes). Prudential norms are based on rationality (e.g one ought to maximize one's expected utility). When members of a society violate the societal norms, they may be punished.

Many social scientists have studied why norms are adhered to. Some of the reasons for norm adherence include:

- fear of authority or power
- rational appeal of the norms
- emotions such as shame, guilt and embarrassment that arise because of non-adherence.
- willingness to follow the crowd

Elster [1] categorizes norms into consumption norms (e.g. manners of dress), behaviour norms (e.g. the norm against cannibalism), norms of reciprocity (e.g. gift-giving norms), norms of cooperation (e.g. voting and tax compliance) etc.

For the purpose of this paper, we focus on social norms because the agents in multi-agent systems have been modelled using ideas borrowed from sociology such as speech act theory and autonomy. Software agents are the proxies for human agents and possess these human-like attributes. Agents acting on behalf of humans (e.g. in virtual worlds) or as independent entities (bots) will need this notion of social norms that regulate their behaviour. Based on the definitions provided by various researchers, we note that the notion of a norm is generally made up of the following two aspects.

- Normative expectation of a behavioural regularity: There is a general agreement within the society that a behaviour is expected on the part of an agent (or actor) by others in a society, in a given circumstance.
- A norm spreading factor : Examples of norm spreading factors include the notion of advice from powerful leaders and the sanctioning mechanism. When an agent does not follow the norm, it could be subjected to a sanction. The sanction could include monetary or physical punishment in the real world which can trigger emotions (embarrassment, guilt etc.) or direct loss of utility. Other kind of sanctions could include agents not being willing to interact with an agent that violated the norm or the decrease of its reputation

score. Other norm spreading factors include imitation and learning on the part of an agent.

It should be noted that researchers are divided on what the differences between a norm and a convention are. Our belief is that convention is a common expectation amongst (most) others that an agent adopts a particular action or behaviour (e.g. the convention in ancient Rome was to drive on the left). In this paper we do not distinguish conventions from norms. Both of them have been incorporated under the umbrella of *norms*.

2.1 Normative multi-agent systems

Research on norms in multi-agent systems is about two decades old. [9–11]. Norms in multi-agent systems are treated as constraints on behaviour, goals to be achieved or as obligations [12].

The definition of normative multi-agent systems as described by the researchers involved in the NorMAS 2007 workshop is as follows [13]. *A normative multi-agent system is a multi-agent system organized by means of mechanism to represent, communicate, distribute, detect, create, modify and enforce norms, and mechanisms to deliberate about norms and detect norm violation and fulfillment.*

The research in normative multi-agent systems can be categorized into two branches. The first branch focuses on normative system architectures, norm representations, norm adherence and the associated punitive or incentive measures. Lopez et al. [14] have designed an architecture for normative BDI agents and [15] have proposed a distributed architecture for normative agents. Some researchers are working on using deontic logic to define and represent norms [15, 16]. Several researchers have worked on mechanisms for norm compliance and enforcement [17–19]. A recent development is the research on emotion based mechanisms for norm enforcement [20, 21]. Conte and Castelfranchi [22] have worked on an integrated view of norms. Their work tries to bridge the gap between the prescriptive view of norms and the emergence of conventions from mere regularities using cognitive abilities of an agent. For a comparison of normative architectures refer to Neumann's article [23].

The second branch of research is related to emergence of norms. Neumann has presented a case study of four research works on the simulation models of norms from the perspective of foundations of social theory [24]. In this work, the four papers were investigated in detail for identifying three methodological core problems which are norm transmission, norm transformation and the function of the norm. The first two problems correspond to the causal aspect of the norm (i.e. what causes the norm to spread). The last problem deals with the purpose of the norm. The author concludes that no model has been able to fully explain both the causal and functional reasons behind norm emergence, however, the current trend is towards bridging this gap.

3 Phases of norm life-cycle

Researchers interested in norms have experimented with several mechanisms associated with norms. Firstly we identify four phases of the norm life-cycle. Secondly, we categorize the simulation mechanisms into 10 categories and have assigned each category to a particular phase of the norm life-cycle.

In the body of research literature on social norms there isn't a unified view on how norms are created and spread in a society. Several researchers have proposed models of norms [1, 7, 25–28]. In this paper we refer to four important phases of norm life-cycle which are norm creation, spreading, enforcement and emergence. Even though there hasn't been any agreement on these phases by social researchers, we use these four phases as they broadly capture the processes associated with the norm life-cycle. Figure 1 shows the four phases of norm life-cycle (in the left) and the categories of mechanisms (in the right).

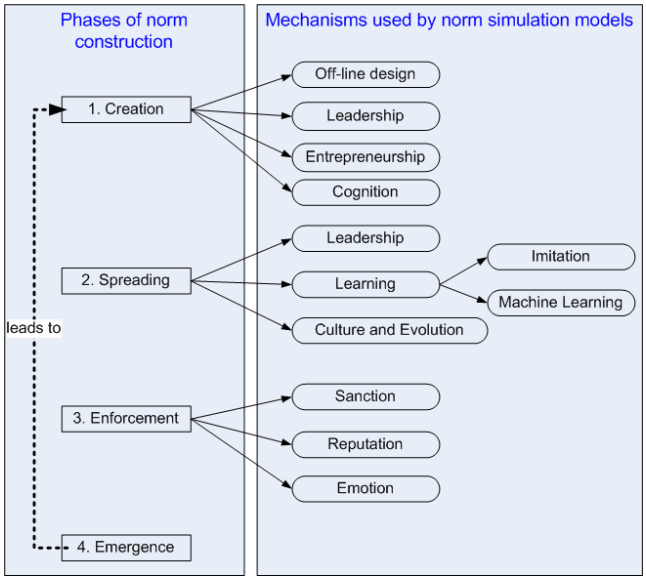

Fig. 1: Phases of norm life-cycle and categories of simulation models

The first phase of the life-cycle model is that of norm creation. A norm can be created by a designer of the system or a powerful leader. The designer and leadership approaches are top-down authoritarian approaches. The other approach for norm creation is the entrepreneurial approach where an agent might come up with a norm and can recommend the norm to other agents. These norms when created are the "proposed norms". Once such a proposed norm is created by a designer, leader or entrepreuneur, it spreads through the society by one of the spreading mechanisms such as advice about a norm by powerful members of the society, imitation, learning on the part of the agents, cultural inheritance

and evolutionary inheritance. Thus, norm spreading forms the second phase of the norm life-cycle. When norms have spread and are internalized, agents may expect other agents to follow the norm that they have subscribed to and may sanction those agents that do not follow the norm. The third phase of the life-cycle is the enforcement of norms where norm violators may be punished, their utility might be reduced, their reputation impacted or emotions such as shame and guilt being stirred which help in the regulation of normative behaviour. The fourth phase is the norm emergence phase. A norm can be said to have emerged if it has spread (i.e. it is followed by a considerable proportion of an agent society and this fact is recognized by most agents). Another aspect of norm emergence is that a norm can emerge without being explicitly created. Norms can emerge in a bottom-up way. One or more cognitive agents, based on interactions in an agent society can infer what the norms of the society are. We can say that these agents derived a "proposed norm" based on their cognitive ability (creation phase) and then helped in the emergence of that norm (emergence phase). These cognitive agents can also come up with an alternative norm that spreads and emerges in a society and hence can replace an existing norm. This feedback loop is represented as a dashed line in 1. It should be noted that not all simulation based research on norms have considered all the 4 phases.

The life-cycle that we have presented is similar to Finnemore and Sikkink's model [27]. They have proposed three stages for the norm life cycle. The first stage is the norm emergence stage which is characterized by the persuasion by some norm entrepreneurs or norm innovators. Norm entrepreneurs are the innovators who think about new ideas/norms in a society. Norm entrepreneurs attempt to convince a critical mass of norm leaders to embrace new norms. The second stage is characterized by the dynamics of imitation as the norm leaders attempt to socialize with other people whom they might have influence over, so they might become followers. The third stage is the norm cascade stage where the followers take up the norm for reasons such as pressure to conform. As the reader may observe, this model is a subset of the life-cycle model that we have proposed. Also, this model caters for the entrepreneurial approach for norm creation and the imitation approach for norm spreading. But, in our approach, more mechanisms are brought under each of the phases.

3.1 Norm creation

Norm creation in multi-agent system refers to the mechanism by which an agent in the society comes to know what the norm of the society is. There are three approaches that simulation works have used and they are a) a designer specified norms (off-line design) b) a norm-leader specified norms and c) a norm-entrepreneur considers that a norm is good for the society.

Off-line design approach - In this approach, norms are designed off-line, and hardwired into agents. Walker and Wooldridge [29] note the following about the off-line design of norms. *"The off line design of norms will often be simpler*

to implement and might present the designer with a greater degree of control over system functionality. However, there are a number of disadvantages with this approach. First, it is not always the case that all the characteristics of a system are known at design time; this is most obviously true of open systems Secondly, in complex system the goals of agents might be constantly changing. To keep reprogramming agents in such circumstances would be costly and inefficient. Finally, the more complex the a system becomes, the less likely it is that system designers will be able to design effective social laws".

Some researchers have used this approach to compare the performance of a normative system with a non normative one [30].This approach is only suitable for top-down norm prescription that is a characteristic of closed and centralized institutions.

Leadership approach - In this approach, some powerful agents in the society (the norm-leaders) come up with a norm. The leader can provide these norms to the follower agents [10, 31].

Entrepreneurial approach - In agent societies, there might be some norm-entrepreneurs who come up with a norm. When an agent comes up with a new norm it tries to convince other agents [32].

Cognitive approach - One or more cognitive agents in a society can come up with norms based on the deliberative processes that they employ [33]. In this approach the agents have the cognitive ability to recognize what the norms of a society are based on the observations of interactions. It should be noted that the norm inferred by each agent might be different (which is based on the observations that an agent has made). Thus, an agent in this model creates its notion of what the norm is based on inference.

3.2 Norm spreading

Norm spreading relates to the distribution of a norm among a group. Mirriam-Webster's dictionary [34] defines spreading as *to become dispersed, distributed, or scattered or to become known or disseminated*. There are several mechanisms that help in spreading the norms such as leadership, imitation , machine learning, cultural and evolutionary mechanisms. These mechanisms are discussed in detail in the next section.

3.3 Norm enforcement

Norm enforcement refers to the process by which norm violators are discouraged through some form of sanctioning. A widely used sanctioning mechanism is the punishment of a norm violator (e.g. monetary punishment which reduces the

agents fitness or a punishment that invokes emotions such as guilt and embarrassment). Reputation mechanisms have also been used as sanctions where an agent is black-listed for not following a norm. The process of enforcement helps to sustain norms in a society. Some researchers have considered enforcement as a part of the spreading mechanism [19].

3.4 Norm emergence

We define norm emergence to be reaching some significant threshold in the extent of the spread of a norm. For example, a society is said to have a norm of gift exchange at Christmas if more than $x\%$ of the population follows such a practice. The value of x varies from society to society and from one kind of norm to another. The value of x has varied from 35 to 100 across different simulation based studies of norms.

Simulation research on norms has employed two approaches to norm emergence. One approach is that an agent comes to know about a norm through mechanisms such as leadership [31, 35] or through imitation [3] and when it accepts the norm it contributes to norm spreading and emergence. The other way is that a cognitive agent could generate a personal norm based on observation [36]. Additionally many such cognitive agents in the society could generate similar personal norms and for an external observer it might seem that a norm has emerged in a society. Also, cognitive agents could communicate norms and verify norms. The later bottom-up approach where micro interactions between agents that lead to the macro effect of establishing a norm is more interesting than the leadership and imitation based approaches.

4 Categorization of simulation works of norm creation, spreading and emergence

In this section we categorize simulation work on norms into eight main categories (shown in figure 2). Each category corresponds to a particular mechanism (e.g. sanctioning mechanism, reputation mechanism). For each of these categories we provide a brief description and discuss a few key papers. It should be noted that some papers have made use of mechanisms that fall under more than one category.

4.1 Social power

Social power plays an important role in societies in establishing order and enabling smoother functioning. Several researchers in normative multi-agent systems have focused on the notion of power [37–39] such as institutional power. Lopez in her thesis on social power and norms notes that powers of an agent are expressed through its abilities to change the beliefs, the motivations, and the goals of other agents in such a way that its goals can be satisfied [40].

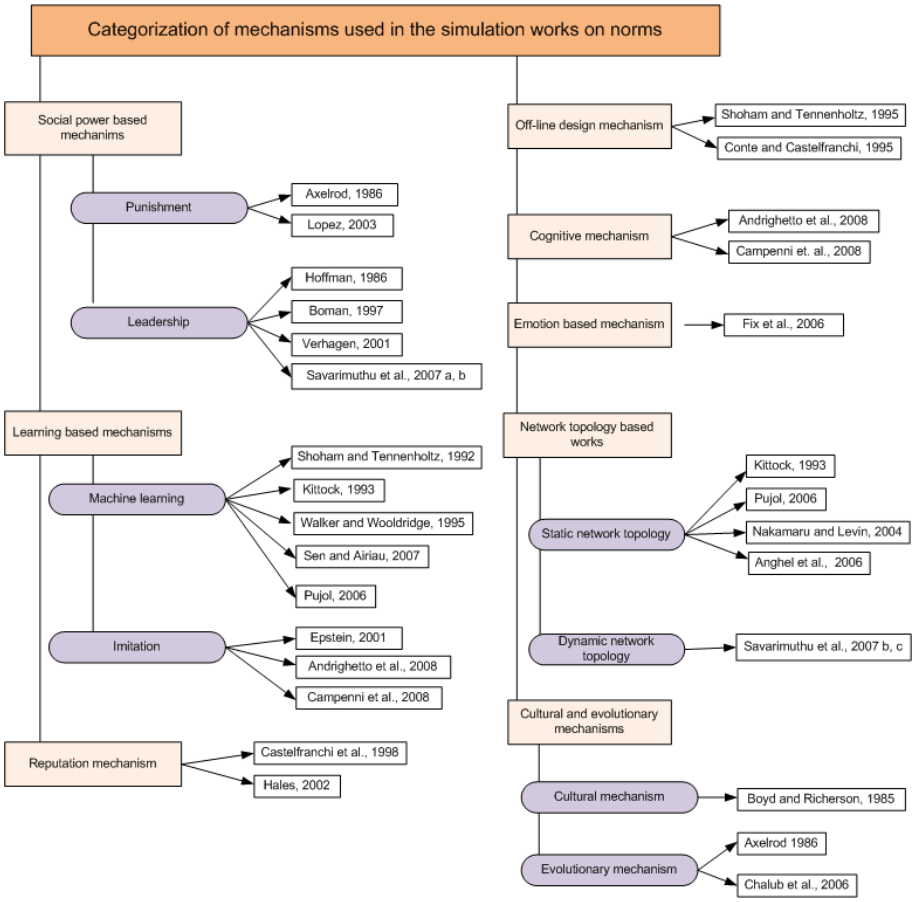

Fig. 2: Categorization of simulation models

Sources of power could motivate, encourage or coerce their followers to take up a particular norm (leadership approach) or force them to adopt a particular norm based on sanctions (punishment approach). Researchers have experimented with both types of social power approaches for norm spreading and enforcement.

Leadership mechanism - Leadership mechanisms are based on the notion that there are certain leaders in the society. These leaders provide advice to the agents in the society. The follower agents seek the leaders advice when deciding about a norm. Verhagen [31] has used the concept of normative advice (advice from the leader of a society) as one of the mechanisms for spreading and internalizing norms in an agent society. However, this centralized approach might not work well in open, flexible and dynamic societies. Savarimuthu et al. [35] extended Verhagen's model by adopting a distributed mechanism for norm emergence. In their mechanism, there could be several normative advi-

sors or role models whom other agents can request for advice. In this model, an agent can be a leader for some agents while that agent itself can be a follower of some other agent. Hoffmann [32] has experimented with the notion of norm entrepreneurs who think of a norm that might be beneficial to the society. His experiments explore the entrepreneurial norm dynamics and provide some initial evidence for Finnemore and Sikkink's norm life cycle model [27].

Sanction mechanism - Even though the models discussed above are based on the notion of power and leadership, they do not include the notion of sanctioning agents that do not follow the norm specified by a norm leader. Several works on norms have used the notion of social power to inflict sanctions on agents that do not follow a norm [17,19,41]. In his well known work [19], Axelrod has shown how a meta-norm that defections should be punished can bring about the norm of cooperation. Lopez et al. [17] have considered punishments and rewards in their model. Their framework models agents with different personalities (social, pressured, opportunistic, greedy, fearful, rebellious). A proper account of the cost of punishment has not been considered in both these works. While Axelrod's work does not consider cost of punishment on the part of the sanctioning agent, Lopez's work assumes that a third party somehow bears this cost. Flentge et al. [41] have shown how an agent comes to acquire a possession norm. They have noted that sanctions help in the establishment of the possession norm if the sanctioning costs are low or when there is no cost for sanctioning.

4.2 Reputation mechanism

Reputation refers to the positive or negative opinion about a person or agent based on their interactions with others in the society. Researchers [42, 43] have addressed how reputation models are beneficial in sustaining norms in an agent society. They have experimented with the effect of the normative reputation on the compliance costs of the norm. They have shown that the normative reputation of agents of the society helps in redistributing the costs of norm compliance to both the agents that follow the norms as well as those who do not follow the norms.

4.3 Imitation mechanism

The philosophy behind an imitation mechanism is *When in Rome, do as Romans do* [3]. These models are characterized by agents mimicking the behaviour of what the majority of the agents do in a given agent society (following the crowd). Epstein's main argument for an imitation mechanism is that individual thought (i.e. the amount of computing needed by a agent to infer what the norm is) is inversely related to the strength of a social norm [3]. This implies that when a norm becomes entrenched the agent can follow it without much thought. Epstein has demonstrated this in the context of a driving scenario in which agents

can observe each other's driving preference (left or right) based on a certain observation radius r. If the agent sees more agents driving on the right within the observation radius,it changes to the right. When a norm is established, the radius tends to move towards one. Other researchers have also experimented with imitation models [36,40,44]. This might be a good mechanism when agents want to avoid the cost of thinking about what the norm of the society is. An agent using the imitation model is not involved in the creation of the norm, it is just a part of the norm spreading effort. Though simple, the model can only account for a way to spread the norm (which is blindly following it). It has been noted that imitation approach cannot bring about the co-existence of multiple norms in a society [45, 46]. Also, it is debatable if imitation-based behaviour (solely) really leads to norms as there is no notion of common expectation.

4.4 Off-line design approach

Off-line design models are characterized by the agents of the society possessing explicit knowledge of the norm. The intention of the designer specified approach is to see how the society performs when the whole society possesses a norm. One of the well-known works on norms specified by the designer is Shoham and Tennenholtz [9]. They have experimented with norms associated with traffic. Several other researchers [29, 43, 47] have experimented with an off-line design approach borrowing the basic experimental set-up proposed by Conte and Castelfranchi [47]. Conte and Castelfranchi have shown using their simulation experiments what the function of a norm is in the context of agents finding food in a grid environment characterized by simple rules for movement and food collection. They have compared the utilitarian strategy with the normative strategy. They have shown that norms reduce the aggression level of the agent (i.e. when a finder-keeper norm is followed) and also increase the average strength of an agent.

4.5 Works based on machine learning

Several researchers have experimented with agents finding a norm based on learning on the part of an agent [2, 29, 48]. Shoham and Tennenholtz have used a mechanism called co-learning which is a simple reinforcement learning mechanism. They have used the "Highest Cumulative Reward (HCR)" rule to update an agent's strategy when playing a simple coordination game and a cooperation game (prisoner's dilemma). According to this rule, an agent chooses the strategy that has yielded the highest reward in the past m iterations. The history of the strategies chosen and the rewards for each strategy is stored in a memory of a certain size (which can be varied). Walker and Wooldridge's experimental model [29] is based on the work done by Conte and Castelfranchi [47] where agents move about a grid in search of food. They have experimented with 16 mechanisms for norm emergence. Their model used two parameters, the majority and the strategic update function. Each of these parameters can be varied

with four values. 16 experiments were based on size of the majority (simple, double, quadruple, dynamic) and the nature of the update function (using majority rule, memory restart, communication type and communication on success). Sen and Airiau [48] have proposed a mechanism for the emergence of norms through social learning. They have experimented with three reinforcement learning algorithms and the agents learn norms based on private local interactions. They have observed that when the population size is bigger the norm convergence is slower and larger the set of possible action states the slower is the convergence. They have also studied the influence of adding agents with a particular action state to a pool of existing agents as well as norm emergence in isolated sub-populations.

Learning mechanisms employ a particular algorithm to identify a strategy that maximizes an agent's utility and the chosen strategy is declared as the norm. Since all agents in the society make use of the same algorithm, the society stabilizes to a uniform norm. Agents using this approach cannot distinguish between a strategy and a norm. The agents do not have a notion of normative expectation (i.e. others expect certain behaviour on the part of an agent) associated with a norm.

4.6 Cognitive approach

Researchers involved in the EMIL project [33] are working on a cognitive architecture for norm emergence. There have been some attempts to explore how the mental capacities of agents play a role in the emergence of norms.

EMIL project aims to deliver a simulation-based theory of norm-innovation, where norm-innovation is defined as a 2-way dynamics of inter-agent process and intra-agent process. The inter-agent process results in the emergence of norms where the micro interactions produce macro behaviour (norms). The intra-agent process refers to what goes inside an agent's mind so that they can recognize what the norms of the society are. This approach is different from the learning models as the agents in the cognitive approach are autonomous and have the capability to examine interactions between agents and are able to recognize what the norms could be. The agents in this model need not necessarily be utility maximizing like the ones in the learning models. The agents in the model will have the ability to filter external requests that affect normative decisions and will also be able to communicate norms with other agents. Agents just employing learning algorithms lack these capabilities.

Andrighetto et al. [36] have demonstrated how the norm recognition module of the EMIL-A platform answers the question "how does a agent come to know of what a norm is". In particular they have experimented with an imitation approach versus the norm recognition approach that they have come up with. The norm recognition module consists of two constructs, the normative board and a module for storing different types of modals for norms. Each modal represents a type of message that is exchanged between agents (e.g. deontics modal refers to partitioning situations as either acceptable or unacceptable). The normative board consists of normative beliefs and normative goals. They have shown that norm recognizers perform better than social conformers (imitating agents) by

the fact that the recognizers were able to identify a pool of potential norms while the imitators generated only one type of norm.

The limitation of this approach is that agents just observe actions performed by other agents. In practice they should be able to learn from their own experience as well. Perhaps, their own experience can be given a higher weight. At present, agents in their model do not have the capability of violating the norms and hence there are no costs associated with sanctions. The authors note this can be a potential extension.

4.7 Emotion based works

Based on the previous work done by Scheve et al. [20], Fix et al. [21] discuss the micro-macro linkage between emotions at the micro-level and the norm enforcement at the macro-level. The authors argue that emotions have a norm regulatory function in agent societies. An agent observing a deviation of a norm might generate emotions such as contempt or disgust which can be the motivation behind sanctions. Those agents that are sanctioned might generate emotions such as shame, guilt or embarrassment which might lead to norm internalization. The authors have used a Petri net model [49] to capture the micro-macro linkage. It should be noted that the proposed model has not been implemented in the context of a simulation experiment. Staller and Petta [50] have extended Conte et al.'s experimental set up by including emotion based strategies.

4.8 Works using network topologies

Social networks are important for norm spreading and emergence because in the real world, people are not related to each other by chance. They are related to each other through the social groups that they are in, such as the work group, church group, ethnic group and hobby group. Information tends to percolate among the members of the group through interactions. Also, people seek advice from a close group of friends and hence information gets transmitted between the members of the social network.

In most simulation works, the treatment of norms has been mostly in the context of an agent society where the agents interact with all the other agents in the society [10, 31] in a random fashion. Few researchers have considered the actual topologies of the social network for norm emergence [44]. We believe such an approach is important for the study of norm spreading and emergence as networks provide the topology and the infrastructure on which the norms can be exchanged. Researchers have studied different kinds of network topologies and their applications in the real world (a overview of different topologies is given by Mitchell [51]). These application areas include opinion dynamics [52] and the spread of diseases [53]. Researchers in normative multi-agent systems have started to look at the role of network topologies [44, 54–56]. Network topologies have also been explored by other multi-agent system researchers in other contexts such as reputation management [57, 58].

Research that has considered network topologies can be categorized into static and dynamic network topology approaches. In the static approach, the network topology is fixed. In the dynamic topology approach, the underlying network can change when the simulation experiments are conducted.

Works using a static network topology - Kittock was the first to experiment with the role of network topology in convention emergence [54]. He noted that the choice of the global structure has a profound effect on the evolution of the system. Pujol's PhD thesis [44] dealt with the emergence of conventions on top of social structures. He used the HCR mechanism proposed by Shoham and Tennenholtz [2] to test norm emergence in connected, random, small world and scale-free networks. He also demonstrated that the structure of the network is crucial for norm emergence. Nakamaru and Levin [46] studied how two related norms evolve in networked environments. Anghel et al. [59] investigated the effects of inter-agent communication across a network in the context of playing minority game. They have shown that a scale-free leadership structure emerges on top of a random network.

Dynamic topology works - Very few researchers have investigated the role of dynamic network topologies on norm spreading and emergence. Savarimuthu et al. [55] used Gonzalez et al.'s model [60] to create dynamic network topologies. Gonzalez et al. have developed a model for constructing dynamically changing networks. They have used the concept of agents (or particles) colliding in an abstract social space to construct evolving networks. Savarimuthu et al. [55] have created dynamic network topologies using Gonzalez's model on which they test their role model agent-based leadership mechanism. They have shown how different types of norms emerge when societies with different norms for the same context (playing the Ultimatum game [61]) are brought together. In particular, they have shown that under certain conditions norms can co-exist in an agent society.

4.9 Cultural and evolutionary mechanisms

Researchers have also proposed other mechanisms for norm spreading and emergence. These include cultural and evolutionary models [62, 63]. Boyd and Richerson [62] have proposed that norms can be propagated through cultural transmission. According to them, there are three ways by which a social norm can be propagated from one member of the society to another. They are

- Vertical transmission (from parents to offspring)
- Oblique transmission (from a leader of a society to the followers)
- Horizontal transmission (from peer to peer interactions)

Of these three kinds of norm transmission mechanisms, vertical and oblique transmissions can be thought of as leadership mechanisms in which a powerful

superior convinces the followers to adopt a norm. The horizontal transmission is a peer-to-peer mechanism where agents learn from day-to-day interactions from other peers. Few researchers have used this idea to experiment with norm spreading [31, 64].

A few researchers have experimented with norm spreading based on evolution where the offsprings inherit the behaviour of the parents. One well known work in this category is Axelrod's [19]. Few other researchers have also experimented with evolutionary models for norm spreading [56, 63]. Chalub et al. [63] have experimented on how norms might spread in different societies (e.g. an archipelago of islands). Agents in an island are fully connected to each other. Each agent plays the donor-receiver game once with all other agents in the island. Then an agent reproduces by choosing a connected agent at random and comparing the payoff. If its payoff is higher than the other agent, then the other agent inherits the strategy of the winning player. Each island has a Gross Domestic Product (GDP) which is a normalized average payoff of the entire island. Islands compete against each other. There are times of war and peace. During peace times, the norms of the islands do not change. When the islands are at war, they play the Hawk and Dove [65] game. The losers change their norm based on a probabilistic norm update rule. The authors note that a meta-norm is established at the end of each run. One limitation of this approach is that they assume that norms have somehow been internalized by a parent/propagator.

Table 1 shows the mechanisms used by the various simulation works on norms corresponding to each phase of the norm life-cycle. It should be noted that not all phases of norm life-cycle have been taken into account by most works.

5 Conclusions

This paper has made two contributions to normative multi-agent system field in the context of simulation of norms. Firstly, a four phase model of the norm life-cycle was proposed. Secondly, various norm-based simulation works were categorized based on the mechanisms employed by each of the works. In the future, we intend to elaborate the research that has been carried out using each of the mechanisms discussed in this paper and also compare their strengths and weaknesses. We will also compare the simulation works based on the agent characteristics employed in each of the works. We also intend to discuss the research issues that need to be addressed.

References

1. Elster, J.: Social norms and economic theory. The Journal of Economic Perspectives **3**(4) (1989) 99–117
2. Shoham, Y., Tennenholtz, M.: Emergent conventions in multi-agent systems: Initial experimental results and observations (preliminary report). In: KR. (1992) 225–231
3. Epstein, J.M.: Learning to be thoughtless: Social norms and individual computation. Comput. Econ. **18**(1) (2001) 9–24

4. : Second life. http://secondlife.com/
5. Habermas, J.: The Theory of Communicative Action : Reason and the Rationalization of Society. Volume 1. Beacon Press (1985)
6. Ullmann-Margalit, E.: The Emergence of Norms. Clarendon Press (1977)
7. Coleman, J.: Foundations of Social Theory. Belknap Press (August 1990)
8. Tuomela, R.: The Importance of Us: A Philosophical Study of Basic Social Notions. Stanford Series in Philosophy, Stanford University Press (1995)
9. Shoham, Y., Tennenholtz, M.: On social laws for artificial agent societies: Off-line design. Artificial Intelligence **73**(1-2) (1995) 231–252
10. Boman, M.: Norms in artificial decision making. Artificial Intelligence and Law **7**(1) (1999) 17–35
11. Conte, R., Falcone, R., Sartor, G.: Agents and norms: How to fill the gap? Artificial Intelligence and Law **7**(1) (1999) 1–15
12. Castelfranchi, C., Conte, R.: Cognitive and social action. UCL Press, London (1995)
13. Boella, G., Torre, L., Verhagen, H.: Introduction to the special issue on normative multiagent systems. Autonomous Agents and Multi-Agent Systems **17**(1) (2008) 1–10
14. López y López, F., Márquez, A.A.: An architecture for autonomous normative agents. In: Fifth Mexican International Conference in Computer Science (ENC'04), Los Alamitos, CA, USA, IEEE Computer Society (2004) 96–103
15. Boella, G., van der Torre, L.: An architecture of a normative system: counts-as conditionals, obligations and permissions. In: AAMAS, New York, NY, USA, ACM Press (2006) 229–231
16. García-Camino, A., Rodríguez-Aguilar, J.A., Sierra, C., Vasconcelos, W.: Norm-oriented programming of electronic institutions. In: Proceedings of the fifth international joint conference on autonomous agents and multiagent systems, AAMAS, New York, NY, USA, ACM Press (2006) 670–672
17. López y López, F., Luck, M., d'Inverno, M.: Constraining autonomy through norms. In: Proceedings of The First International Joint Conference on Autonomous Agents and Multi Agent Systems AAMAS'02. (2002) 674–681
18. Aldewereld, H., Dignum, F., García-Camino, A., Noriega, P., Rodríguez-Aguilar, J.A., Sierra, C.: Operationalisation of norms for usage in electronic institutions. In: AAMAS, New York, NY, USA, ACM Press (2006) 223–225
19. Axelrod, R.: An evolutionary approach to norms. The American Political Science Review **80**(4) (1986) 1095–1111
20. Scheve, C., Moldt, D., Fix, J., Luede, R.: My agents love to conform: Norms and emotion in the micro-macro link. Comput. Math. Organ. Theory **12**(2-3) (2006) 81–100
21. Fix, J., von Scheve, C., Moldt, D.: Emotion-based norm enforcement and maintenance in multi-agent systems: foundations and petri net modeling. In: AAMAS. (2006) 105–107
22. Conte, R., Castelfranchi, C.: From conventions to prescriptions - towards an integrated view of norms . Artif. Intell. Law **7**(4) (1999) 323–340
23. Neumann, M.: A classification of normative architectures. In: WCSS-08 Proceedings, not known (2008) not known
24. Neumann, M.: Homo socionicus: a case study of simulation models of norms. Journal of Artificial Societies and Social Simulation **11**(4) (2008) 6
25. Opp, K.D.: How do norms emerge? An outline of a theory. Mind and Society **2**(1) (2001) 101–128

26. Horne, C.: Sociological perspectives on the emergence of norms. Social Norms (Hechter, M. and Opp, KD, eds) (2001) 3–34
27. Finnemore, M., Sikkink, K.: International Norm Dynamics and Political Change. International Organization **52**(04) (2005) 887–917
28. Bicchieri, C.: The Grammar of Society - The Nature and Dynamics of Social Norms. Cambridge University Press (2006)
29. Walker, A., Wooldridge, M.: Understanding the emergence of conventions in multi-agent systems. In Lesser, V., ed.: Proceedings of the First International Conference on Multi–Agent Systems, San Francisco, CA, MIT Press (1995) 384–389
30. Conte, R., Castelfranchi, C.: Norms as mental objects - from normative beliefs to normative goals. In: MAAMAW. (1993) 186–196
31. Verhagen, H.: Norm Autonomous Agents. PhD thesis, Department of Computer Science, Stockholm University (2000)
32. Hoffmann, M.: Entrepreneurs and Norm Dynamics: An Agent-Based Model of the Norm Life Cycle. Technical report, Department of Political Science and International Relations, University of Delaware, USA (2003)
33. Andrighetto, G., Conte, R., Turrini, P., Paolucci, M.: Emergence in the loop: Simulating the two way dynamics of norm innovation. In Boella, G., van der Torre, L., Verhagen, H., eds.: Normative Multi-agent Systems. Number 07122 in Dagstuhl Seminar Proceedings, Internationales Begegnungs- und Forschungszentrum fuer Informatik (IBFI), Schloss Dagstuhl, Germany (2007)
34. : Spreading - definition from the merriam-webster online dictionary
35. Savarimuthu, B.T.R., Purvis, M.A., Cranefield, S., Purvis, M.K.: How do norms emerge in multi-agent societies? mechanisms design. (2007)
36. Andrighetto, G., Campenni, M., Cecconi, F., Conte, R.: How agents find out norms: A simulation based model of norm innovation. In: Not known, not known (2008) not known
37. Castelfranchi., C., Cesta, A., Miceli, M.: Dependence relations among autonomous agents. In: Decentralized A.I.-3. Elsevier, Amsterdam (1992)
38. Jones, A.J.I., Sergot, M.J.: A formal characterisation of institutionalised power. Logic Journal of the IGPL **4**(3) (1996) 427–443
39. Castelranchi, C.: C. Castelfranchi. All I understand about power (and something more). Technical report, ALFEBIITE Project, London (2000)
40. López y López, F.: Social Powers and Norms: Impact on Agent Behaviour. PhD thesis, University of Southampton, United Kingdom (2003)
41. Flentge, F., Polani, D., Uthmann, T.: Modelling the emergence of possession norms using memes. Journal of Artificial Societies and Social Simulation **4** (2001)
42. Castelfranchi, C., Conte, R., Paolucci, M.: Normative reputation and the costs of compliance. Journal of Artificial Societies and Social Simulation **vol. 1, no. 3** (1998)
43. Hales, D.: Group reputation supports beneficent norms. Journal of Artificial Societies and Social Simulation **5** (2002)
44. Pujol, J.M.: Structure in Artificial Societies. PhD thesis, Software Department, Universitat Politénica de Catalunya (2006)
45. Campenni, M., Andrighetto, G., Cecconi, F., Conte, R.: Normal = normative? the role of intelligent agents in norm innovation. In: Not known, not known (2008) not known
46. Nakamaru, M., Levin, S.A.: Spread of two linked social norms on complex interaction networks. Journal of Theoretical Biology **230**(1) (September 2004) 57–64

47. Conte, R., Castelfranchi, C.: Understanding the effects of norms in social groups through simulation. In: Artificial societies: the computer simulation of social life. UCL Press, London (1995)
48. Sen, S., Airiau, S.: Emergence of norms through social learning. In: Proceedings of Twentieth International Joint Conference on Artificial Intelligence (IJCAI), Hyderabad, India, MIT Press (2006) 1507–1512
49. Jensen, K., ed.: Application and Theory of Petri Nets 1992, 13th International Conference, Sheffield, UK, June 22-26, 1992, Proceedings. In Jensen, K., ed.: Application and Theory of Petri Nets. Volume 616 of Lecture Notes in Computer Science., Springer (1992)
50. Staller, A., Petta, P.: Introducing emotions into the computational study of social norms: A first evaluation. J. Artificial Societies and Social Simulation **4**(1) (2001)
51. Mitchell, M.: Complex systems: Network thinking. Artificial Intelligence **170**(18) (2006) 1194–1212
52. Fortunato, S.: Damage spreading and opinion dynamics on scale free networks (2004)
53. Cohen, R., Havlin, S., ben Avraham, D.: Efficient immunization strategies for computer networks and populations. Physical Review Letters **91** (2003) 247901
54. Kittock, J.E.: Emergent conventions and the structure of multi-agent systems. In Nadel, L., Stein, D.L., eds.: 1993 Lectures in Complex Systems. Addison-Wesley (1995)
55. Savarimuthu, B.T.R., Cranefield, S., Purvis, M.K., Purvis, M.A.: Norm emergence in agent societies formed by dynamically changing networks. In: IAT '07: Proceedings of the 2007 IEEE/WIC/ACM International Conference on Intelligent Agent Technology, Washington, DC, USA, IEEE Computer Society (2007) 464–470
56. Villatoro, D., Sabater-Mir, J.: Categorizing social norms in a simulated resource gathering society. In: Proceeding of the Advancement of Artificial Intelligence (AAAI) workshop on Coordination, Organization,Institutions and Norms in agent systems (COIN). (2008) not known
57. Pujol, J.M., Sangüesa, R., Delgado, J.: Extracting reputation in multi agent systems by means of social network topology. In: Proceedings of the first international joint conference on autonomous agents and multiagent systems, AAMAS, New York, NY, USA, ACM Press (2002) 467–474
58. Yu, B., Singh, M.P.: Searching social networks. In: Proceedings of the second international joint conference on autonomous agents and multiagent systems, AAMAS, New York, NY, USA, ACM Press (2003) 65–72
59. Anghel, M., Toroczkai, Z., Bassler, K.E., Korniss, G.: Competition-driven network dynamics: Emergence of a scale-free leadership structure and collective efficiency. Physical Review Letters **92**(5) (2004) 0587011–0587014
60. Gonzaléz, M.C., Lind, P.G., Herrmann, H.J.: Networks based on collisions among mobile agents. Physica D **224** (2006) 137–148 e-print: physics/0606023.
61. Slembeck, T.: Reputations and fairness in bargaining - experimental evidence from a repeated ultimatum game with fixed opponents. Experimental 9905001, Economics working paper archive (1999)
62. Boyd, R., Richerson, P.J.: Culture and the evolutionary process. University of Chicago Press, Chicago (1985)
63. Chalub, F., Santos, F., Pacheco, J.: The evolution of norms. Journal of Theoretical Biology **241**(2) (2006) 233 – 240
64. Savarimuthu, B.T.R., Cranefield, S., Purvis, M.A., Purvis, M.P.: Role model based mechanism for norm emergence in artificial agent societies. In: Proceeding of the

AAMAS 2007 workshop on Coordination, Organization, Institutions and Norms in agent systems (COIN). (2007) 1–12
65. Smith, M.J., Price, G.R.: The logic of animal conflict. Nature **246**(5427) (November 1973) 15–18

Simulation works	**Norm creation**	**Norm spreading**	**Norm enforcement**	**Norm emergence**
Axelrod, 1986	-	Evolutionary approach	Sanction	Yes
Shoham and Tennenholtz, 1992	Learning	-	-	Yes
Kittock, 1993	-	Learning, network topology	-	Yes
Conte and Castelfranchi, 1995	Off-line	-	-	-
Walker and Woolridge, 1995	-	Learning	-	Yes
Shoham and Tennenholtz, 1995	Off-line	-	-	-
Castelfranchi et al., 1998	Off-line	-	Reputation	-
Verhagen, 2000	Leadership	Leadership	-	-
Epstein, 2001	-	Imitation	-	Yes
Flentge et al., 2001	-	Cultural transmission	Sanction	Yes
Hales, 2002	Off-line	-	Reputation	-
Hoffmann, 2003	Entrepreneurship	leadership	-	Yes
Lopez et al., 2003	Off-line	-	Sanction and reward	-
Nakamaru and Levin, 2004	Off-line	Network topology	-	Yes
Chalub et al., 2006	-	Evolutionary approach	-	Yes
Fix et al., 2006	-	-	Emotion	-
Pujol, 2006	-	Learning, network topology	-	Yes
Sen and Airiau, 2007	-	Learning	-	Yes
Savarimuthu et al., 2007 b,c	-	Leadership, network topology	-	Yes
Andrighetto et al., 2008, Campenni et al., 2008	Cognition	Imitation	-	Yes

Table 1. Mechanisms employed by simulation works in each phase of the norm life-cycle