Norm learning in multi-agent societies

Bastin Tony Roy Savarimuthu
Department of Information Science
University of Otago
Dunedin, New Zealand
tonyr@infoscience.otago.ac.nz

***Abstract*—In Normative Multi-Agent Systems (NorMAS), researchers have investigated several mechanisms for agents to learn norms. In the context of agents learning norms, the objectives of the paper are three-fold. First, this paper aims at providing an overview of different mechanisms employed by researchers for norm learning. Second, it discusses the contributions of different mechanisms to the three aspects of active learning namely learning by doing, observing and communicating. Third, it compares two normative architectures which have an emphasis on the learning of norms. It also discusses the features that should be considered in future norm learning architectures.**

***Keywords*-norms; learning; agents; mechanisms;**

I. Introduction

Norms have been of interest to researchers in multi-agent systems because they enable cooperation and coordination for software agents. They are also light-weight mechanisms for enabling social control. Agents that know about norms in agent societies do not need to recompute what the norms of the society are and also do not often need to spend time in contemplating actions that are forbidden and obliged as they are aware of these norms. Also, agents that are aware of norms know that violating them will have consequences for them. However, this is true only when the agents know what the norms are. A new agent joining a society may not know what the norms of the society are. This agent will need to be equipped with some mechanism for learning the norms.

Researchers have employed several mechanisms for the learning of norms. They include imitation, normative advise from leaders, machine learning and data-mining. A discussion of these approaches are discussed in Section III. In Section IV we discuss three aspects of active learning namely experiential learning, observational learning and communication-based learning and also discuss the use of these three aspects in the existing works on norm learning. In this section, we also discuss the need for integrating these three aspects in the norm learning framework and demonstrate how this can be achieved in the context of a simple example. In Section V we compare two norm learning architectures and also identify the areas for improvement.

II. Background

Due to multi-disciplinary interest in norms, several definitions for norms exist [2]. Elster notes the following about social norms [9]. *"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"*.

Based on the definitions provided by various researchers, we note that the social practices surrounding the notion of a social norm are the following:

- **The 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 enforcement mechanism**: When an agent does not follow a 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 (e.g. decrease of its reputation score).
- **A norm spreading mechanism**: Examples of norm spreading mechanisms include the notion of advice from powerful leaders, imitation and learning on the part of an agent.

It should be noted that knowing the norms that are applicable in a society is a starting point for an agent to incorporate the above mentioned social practices that relate to norms. An agent may come to know about norms in two ways. First, the norm that is applicable may be explicitly hard-coded in its design through the off-line design of norms (refer to [23] for a discussion on off-line design approaches to norms). Second, an agent can learn norms based on the interactions it has in the society. It is the second approach that we focus in this paper.

A. Normative multi-agent systems

The definition of normative multi-agent systems given by the researchers involved in the NorMAS 2007 workshop is as follows [4]. *A normative multi-agent system is a multi-agent 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.*

Researchers in multi-agent systems have studied how the concept of norms can be applied to artificial agents. 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. Researchers have shown that norms improve cooperation and collaboration [28], [32]. Epstein has shown that norms reduce the amount of computation required to make a decision [11]. However, software agents may tend to deviate from norms due to their autonomy. So, the study of norms has become important to MAS researchers as they can build robust multi-agent systems using the concept of norms and also experiment on how norms may evolve and adapt in response to environmental changes.

Research in normative multi-agent systems can be categorized into two branches. Researchers have worked on both prescriptive (top-down) and emergent (bottom-up) approaches to norms. The first branch focuses on normative system architectures, norm representations, norm adherence and the associated punitive or incentive measures. Several architectures have been proposed for normative agents (refer to [20] for an overview). Researchers have used deontic logic to define and represent norms [15], [34]. Several researchers have worked on mechanisms for norm compliance and enforcement such as sanctioning mechanisms [3] and reputation mechanisms [7].

The second branch of research is related to emergence of norms [12], [27], [28]. In the bottom-up approach, the agents come up with a norm through learning mechanisms [27], [28] and cognitive approaches [2]. This paper contributes to this branch of research by providing an overview of the norm learning mechanisms and comparing two of the architectures that have employed learning mechanisms.

III. Approaches to the learning of norms

Researchers have employed mainly four types of mechanisms for an individual agent to learn norms: imitation, machine learning, data mining and advice-based learning (as shown in Figure 1). The learning mechanisms identified in this paper are extensions to the learning mechanisms discussed in the categorization presented in a previous work [23]. Since the imitation, machine learning and advise-based learning mechanisms (also called as leadership mechanisms) are explained in the previous work, we only provide a brief summary of these three approaches and provide a longer discussion on the data mining approaches.

1) Imitation mechanisms: The philosophy behind an imitation-based learning mechanism is *When in Rome, do as the Romans do* [11]. Models based on imitation are characterised by agents first observing and then mimicking the behaviour of what the majority of the agents do in a given agent society (following the crowd). Epstein's main argument [11] for an imitation mechanism is that individual thought (i.e. the amount of computing needed by an agent to infer what the norm is) is inversely related to the strength of a social norm. This implies that when a norm becomes entrenched the agent can follow it without much thought. An issue for debate is whether imitation-based behaviour (solely) really leads to norms as there is no notion of generalized expectation. Imitation mechanism also does not consider sanctions or rewards thereby focus only on conventions and not norms[1]. The direct utility from conforming to a particular behaviour is not modelled in some cases (i.e. blindly imitating what the crowd does without carefully considering the impact of the actions either for the agent or the society).

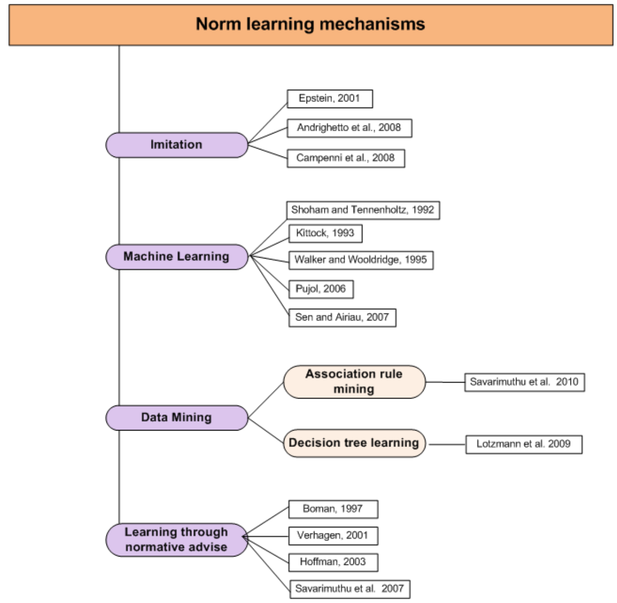

Figure 1. Mechanisms for learning norms in agent societies

2) Works based on machine learning: Several researchers have experimented with agents finding a norm based on learning on the part of an agent when it interacts with other agents in the society by performing some actions [27], [28], [32]. Researchers have used simple reinforcement algorithms for norm learning. The reinforcement learning algorithms 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 stabilises to an uniform norm. Agents using this approach cannot distinguish between a strategy and a norm. These agents accept the strategy that maximizes its utility as its

[1]Many sociologist consider sanctions and/or rewards a core part of the norm.

norm. However, the agents do not have a notion of normative expectation associated with a norm (i.e. when agents expect certain behaviour on the part of other agents). Another weakness is that agents in machine learning approach do not have a mental notion of norms (i.e. the ability to reason about why norms have to be followed, consequences for not following norms) as they are mainly utilitarian agents. These limitations are addressed by the works that employ cognitive approaches where the norms learnt affect an agent's future decision making by influencing its beliefs, intentions and goals [2].

3) Advise-based learning: Boman [5] has used a centralised approach, where agents consult with a normative advisor before they make a choice on actions to perform. Verhagen [30] has extended this notion of normative advice to obtaining normative comments from a centralized normative advisor (e.g. the leader of the society) on an agent's previous choices. Savarimuthu et al. [24] have adopted a distributed approach approach for normative advice. In their mechanism, there could be several normative advisors (called role models) from whom other agents can request advice. Hoffmann [14] has experimented with the notion of norm entrepreneurs who think of a norm that might be beneficial to the society. An entrepreneur can recommend a norm to a certain percentage of the population (e.g. 50%) which leads to varying degrees of establishment of a norm. The models based on advise assume that a powerful authority is present in the society and all agents in the society acknowledge the power of such agents. Both centralised and distributed notions of norm spreading using *power* have been employed. The centralised approach is suitable for closed societies. However, this might not work well for open, flexible and dynamic societies. Distributed approaches for norm spreading and emergence are promising because the computational costs required to spread, monitor and control a norm are distributed across the individual agents.

4) Data mining mechanism: Agents can use a data mining approach to identify norms in agent societies. Agents in open agent societies can learn norms based on what they infer based on their observations of the society. The repository of an agent's observations can then be mined for patterns of behaviour. There has been a proposal of an agent architecture for normative systems to employ data mining for citizens of a country to find information and norms from official documents [29]. However, the work does not describe what types of norms are discovered and also the mechanisms used in the identification of norms.

Savarimuthu et al. [25], [26] have proposed an architecture for norm identification which employs association rule mining, a data mining approach. The architecture makes use signals (sanctions and rewards) as the starting points for norm identification. Mechanisms for identifying two types of norms, prohibition norms and obligations norms have been studied. The details on how an agent identifies a prohibition norm are explained in the context of a public park scenario, where the norm against littering is identified by the agent. The obligation norm inference is explained in the context of a tipping norm in a restaurant scenario. They have demonstrated that an agent using the proposed architecture can dynamically add, remove and modify norms based on mining the interactions that take place between agents. They have shown that agents can identify co-existing a norms. The agents can also identify conditional norms (e.g. identification of normative pre-conditions - conditions that have to be true for the norm to hold).

In the work of Lotzmann et al. [18] an agent learns about a norm by constructing a decision tree of events that occur and learning through the occurrence probabilities of events that take place. For example, an agent participating in a traffic scenario (either as a pedestrian or a car driver), decides based on the probability of events represented by the nodes of a decision tree. Based on these probabilities, a pedestrian agent learns that if it crosses the road other than the pedestrian crossing, it has the high probability of being run over by a car. A car driver learns to stop in the pedestrian crossing area.

Data mining is a promising approach for the identification of some types of norms that can be inferred based on observing the interactions between agents in the society. However, if actions that explicitly signal a sanction or reward are absent or other mechanisms such as reputation are used instead of explicit signals (i.e. reduction in the reputation score of a rogue agent instead of explicit sanctioning), then it is difficult to identify norms.

IV. Aspects of Active Learning of Agents

Hamada et al. [13] note that *active learning is learning with learners involved in the learning process as active partners: meaning they are "doing", "observing" and "communicating" instead of just "listening" as in the traditional learning style*. An actively learning agent can thus learn about norms in the following three ways.

- **Experiential learning** - This is the ability of an agent *learning by doing*. For example, an agent may litter in a park. It may be sanctioned by some other agent(s). Through the sanction experienced as a result of the littering action, the agent can learn about the norm. An agent can learn from its personal experience based on sanctions and rewards.
- **Observational learning** - This is the ability of an agent *learning by observing*. For example, an agent may observe littering agents being sanctioned in a society. Through the observation of the sanction on others, an agent can learn about the norm.
- **Communication-based learning** - This is the ability of an agent *learning by communicating* with other agents. For example, an agent may ask another agent in the park what the norms of the park are and that

agent could communicate the norm to the agent. Norm communication can happen at a peer-to-peer level or from leaders to follower agents.

It can be seen that imitation mechanism is an example of observational learning (i.e. imitating the observed behaviour of other agents). Machine-learning is based on the experiential learning (i.e. interacting with others and learning from the results). The data mining mechanism is based on observational learning (i.e. mining norms based on observed behaviour). Some researchers have considered communication based learning [24], [30], [32].

Table I shows the types of learning used by different research works. It can be noticed that not all the three types of learning are used by research works. However, some recent research works have considered all the three types of learning [18], [26].

Model	**Experiential learning**	**Observational learning**	**Communication-based learning**
Shoham and Tennenholtz, 1992 [28]	Yes	No	No
Kittock, 1993 [17]	Yes	No	No
Walker and Wooldridge, 1995 [32]	Yes	No	Yes
Verhagen, 2001 [30]	No	No	Yes
Epstein, 2001 [11]	No	Yes	No
Hoffmann, 2003 [14]	No	Yes	No
Pujol, 2006 [21]	Yes	No	No
Sen and Airiau, 2007 [27]	Yes	No	No
Savarimuthu et al., 2007 [24]	No	Yes	Yes
Campenní et al., 2008 [6]	Yes	Yes	No
Andrighetto et al., 2010 [1]	Yes	Yes	No
Savarimuthu et al., 2010 [25], [26]	Yes	Yes	Yes
Lotzmann et al., 2009 [18]	Yes	Yes	Yes

Table I
COMPARISON OF THE TYPES OF LEARNING EMPLOYED BY DIFFERENT RESEARCH WORKS (YES - CONSIDERED, NO - NOT CONSIDERED)

A. A simple experiment using all the three types of learning

In this section we demonstrate how the three types of learning can be combined when using machine learning. Research works using machine learning mechanisms have only investigated the experiential learning aspect [21], [27], [31].

Using a simple reinforcement learning mechanism we demonstrate the possibility of combining these three aspects of active learning. Consider the scenario where agents drive either on the left (L) or the right (R) of the road. The payoff matrix for this coordination game is given in Table II. The goal of the learning task is to facilitate all agents to drive either on the right or left. The goal can be achieved through a combination of three aspects of learning. In this work, we will compare three combinations, 1) learning by doing, 2) learning by doing and observing and 3) learning by doing, observing and communicating.

Table II
PAYOFF MATRIX

	L	R
L	1, 1	-1, -1
R	-1, -1	1, 1

Assume that there are 100 agents in the system. In each iteration, the agents randomly interact with one other agent. Based on the outcome of the interaction, the agent learns which action to choose for the next iteration. We used a simple Q-Learning approach [33] to facilitate learning. For a stateless game, the Q-value is calculated using the formula

$$Q(a) = Q(a) + \gamma(R - Q(a)) \quad (1)$$

where Q(a) is the quality (or value) of an agent performing an action a, R is the reward for performing an action and γ, is the learning rate. The agents use the same formula to undertake all the three types of learning.

We have conducted three experiments with the value of γ to be 0.3. In the first experiment, an agent learns only through its experience (i.e. based on the result of their interaction with other agents). In the second experiment, in addition to experiential learning they also observe one other agent's action and learn from the result of that action (experiential + observational learning). In the third experiment, in addition to the set-up of the second experiment, an agent also learns from the experience of one other agent (i.e. by asking about the action performed by the agent and the reward it obtained). It should be noted that the third experiment involves all the three aspects of learning.

Figure 2 shows the results of the three experiments as three lines (results based on the average of 10 runs per experiment). It demonstrates that experiment three that uses a combination of experiential, observational and communication based learning results in the fastest convergence of norms. It is intuitive that an agent that can make use of the three aspects of learning will do better since more information is available to the agent. However, it is interesting to note that not all the three aspects have been considered by many research works as shown in Table I.

B. The need for integrating the three aspects of learning

We note that the future research works on norm learning should consider integrating these three aspects where ever possible. The reasons are outlined below.

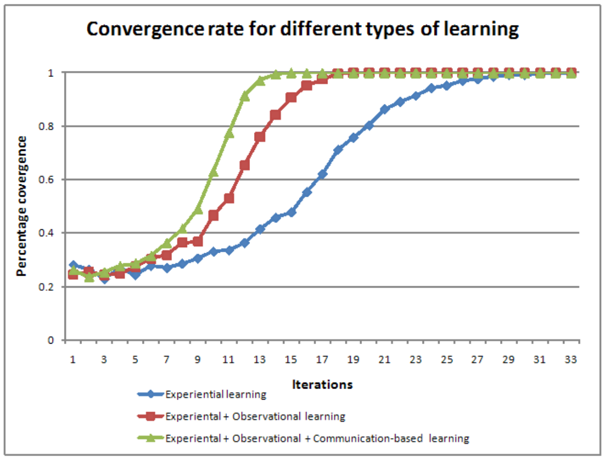

Figure 2. Comparison of convergence rate in three experiments (varying different aspects of learning)

1) One of the drawbacks on the experiential learning of norms in an agent society is that an agent cannot perform all possible actions in order to find out what the norms of the society are. For example, a new agent in a society may not know what the norms are and it may not be desirable to perform all actions to see whether any of those actions result in a sanction by performing it. The state space of actions can be large. Hence, this approach can be computationally expensive. However, if an agent does not actively search for an action that might be sanctioned, but only learns based on receiving a sanction for an action that it performed accidentally, it can use that sanction as a starting point to infer a norm. For example, when it is sanctioned for littering, it can flag the littering action as the potential norm and then check to see in its future interactions with other agents whether littering causes a sanction.
2) Using just the observational learning for learning norms might also cause problems. Assume that agent Z observes agent X punishing Y. Only if Z observes both the action responsible for the sanction and the sanction itself it can learn from the observation. However, if the observer (agent Z) does not know which of the actions agent Y had done in the immediate past had caused this sanction, this approach will not be useful. In this case, it has to either ask another agent for verification or should perform the action under question itself to learn about the norm.
3) Using just the communication learning may be sufficient in regimented societies where norms are prescribed by the organization and in societies where there is no lying. However, in open agent societies it may not be possible to rule out lying. In this case an agent may have to engage in observational learning and/or experiential learning.

In addition to the above mentioned drawbacks on using just one aspect of learning there are some issues with the simple scenario described in Section IV-A.

1) *Limited state space*: Only two actions (turn left or right) are considered. In real life agents may have to know about and also perform different types of actions in a given scenario. Larger state spaces call for the integration of different aspects of learning.
2) *Sanctions/Rewards known ahead of time*: The pay-off matrix specifies what might be the desirable action ahead of time. In open agent societies the action that is being sanctioned may not be known ahead of time and the sanctions/rewards may emerge dynamically and can also change in due course of time. This again calls for the integration of different aspects of learning (e.g. experiencing, observing and communicating norm change).
3) *Limited experiments*: Only three experiments have been presented. However, there is scope for conducting experiments with different combinations (a total of 7 options[2]. The options would depend upon the domain of investigation. In domains where one of the aspects is not possible (say observation is not possible), then the number of options for experimentation will be reduced to 3.

V. Comparison of two architectures that have considered norm learning

So far researchers have dealt with simple scenarios such as the one described in Section III. However, as discussed in Section IV-B there are some issues with these scenarios and also there is a need for combining the three aspects of active learning.

There are several architectures that have been proposed for the study of normative agents [20]. In this study we have chosen two architectures that have an emphasis on the learning of norms, EMIL-A [2] and norm identification [26]. These two independently developed architectures aim to address the question of "how an agent can identify a norm in an open agent society".

A. EMIL architecture (EMIL-A)

Researchers involved in the Emergence In the Loop (EMIL) project [2] have proposed the architecture for norm emergence called EMIL-A that explores how the mental capacities of agents play a role in the emergence of norms. The EMIL project delivers a simulation-based theory of norm innovation, where norm innovation is defined as the

[2]Three options if only one aspect is considered (i.e. doing, observing and communicating), three options when two aspects are considered (doing-observing, observing-communicating, doing-communicating) and one option when three aspects are considered (doing-observing-communicating)

two-way dynamics of an inter-agent process and an 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 on inside an agent's mind so that they can recognise what the norms of the society are.

This approach is different from traditional learning models discussed in Section III, as the agents in the cognitive approach are autonomous and have the capability to interact with other agents, examine interactions between agents and are able to recognise what the norms could be. The agents in this model will have the ability to filter external requests that affect normative decisions and will also be able to communicate norms to other agents. Thus, this cognitive architecture takes into account the three aspects of learning.

The cognitive mechanism employed by this architecture is promising because agents with this type of mechanism have the notion of normative expectation. This mechanism focuses on what goes on inside the mind of an agent to infer norms (i.e. the computational machinery used by the agent to infer norms). Agents infer norms when they join new societies and deliberate about norms. Agents can also suggest a new norm based on their past experience and may bring about norm change.

Andrighetto et al. [1] 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 messages (which the authors call "modals") that can be used to infer norms. The messages that are exchanged between agents can be of five different types (e.g. the deontics modal refers to partitioning situations as either acceptable or unacceptable). The normative board consists of normative beliefs and normative goals, which are modified based on the messages received. They have shown that norm recognisers perform better than social conformers (imitating agents) because the recognisers were able to identify a pool of potential norms while the imitators generated only one type of norm.

EMIL project researchers have demonstrated how their architecture can be used in several scenarios (e.g convention emergence in a traffic scenario and the norm emergence in Wikipedia) [10]. The mechanisms used for norm learning include decision tree learning and reinforcement learning. Genetic algorithm is used for norm innovation (i.e. coming up with a new norm based on applying genetic operators to a norm).

B. Architecture for norm identification

Savarimuthu et al. [25], [26] have proposed an architecture for norm identification. In their architecture, an agent operating in an open agent society has the ability to identify new norms, detect changing norms, and remove norms that do not hold. For this purpose, the agent possesses capabilities of inference based on the observing its local environment. The architecture makes use of signalling as the top level construct to identify both sanctions and rewards. An agent makes use of association rule mining to learn from the observation history. The agent can learn both from personal interactions as well as observed interactions. An agent can also verify whether an identified norm holds by asking another agent. Thus, this architecture also considers all the three aspects of active learning.

Based on the occurrence of these signals, an agent can identify two types of norms, the prohibition norms [25] and the obligation norms [26] by employing association rule mining [8], a data mining approach. Prohibition norms are identified based on the extraction of event sequences that happened in the past that is the reason for a sanction. For example, an agent in Second Life may drop litter in a park. The agent could be sanctioned by another agent. An observer agent (the avatar that is a proxy to the human being) can detect the presence of a norm based on the interactions between agents in its environment. It should be noted that the observer may not have prior knowledge about the norm. It learns about a new norm by recording observed interactions between agents and then applying association rule mining technique on the observed data. Obligation norms are identified based on event sequences that did not happen in the past but were expected to happen which results in a sanction. For example, an agent is expected to tip a waiter in a restaurant. The failure to tip may result in a sanction. Based on observing the events, an agent can infer what the obligation norm could be.

C. Comparison of the two architectures

We note that the architectures proposed by the EMIL project [2] and the norm identification architecture are quite similar in addressing norm learning. They both facilitate all the three aspects of the active learning of norms(i.e. learning by doing, observing and communicating). The overall goal of these two architectures are different. While the norm identification architecture focuses mainly on the question of how an individual agent can learn norms in a society, the EMIL project aims to deliver a socio-cognitive model of norm innovation [10]. The EMIL architecture is more encompassing than the norm identification architecture as takes into account features such as norm immergence and internalization.

Norm identification architecture differs from the EMIL architecture in three ways. First, the norm identification architecture treats "reaction" or "signalling" (positive and negative) to be a top-level construct for identifying potential norms when the norm of a society is being shaped. A sanction in this architecture may not only imply a monetary punishment, but also be an action that could invoke emotions (such as an agent yelling at another might invoke shame or embarrassment on another agent), which can help in norm spreading. Agents can recognize such actions based on their previous experiences. Even though signalling is not explicitly considered as a top level entity for norm learning, several type of message inputs are considered that can point towards a potential norm (e.g. message types such as deontics and evaluation). Second, based on association-rule mining [8], the norm identification architecture makes use of algorithms for norm inference, which can be adapted by an autonomous agent for flexible norm identification. The EMIL project on the other hand makes use of decision-tree mining [18], reinforcement learning and genetic algorithms [10]. Third, using simple examples, how co-existing norms can be identified has been investigated by the researchers involved in the norm identification architecture and they have also pointed to how conflicting norms can be avoided. Even though some abstract examples of co-existing norms have been provided [1], concrete examples on co-existing and conflicting norms seem to be missing in the EMIL project.

D. Additional features that need to be considered

The following features can be considered in the future norm learning architectures.

1) Other forms of norm learning: Other approaches to norm learning such as Partially Observable Markov Decision Processes (POMDPs) [16] and inductive logic programming [19] can be considered for norm learning for individual agents.

2) Emergence of signals (sanctions or rewards): Many current works assume that a sanctioning agent knows a priori the action that should be sanctioned. Though this may hold for norm leader or entrepreneur agents that come up with norms, in some scenarios, the action that is sanctioned may not be known even to the potential sanctioning agent ahead of time. The sanction might emerge depending upon the environmental dynamics. For example, an agent might not sanction if it sees one agent littering. But, when it sees n agents littering the park, it might start punishing, because that action has lowered its utility beyond a certain threshold (an internal utility function). In this scenario, an agent can use a learning algorithm (e.g. Q-Learning [33]) to identify an action that lowers its utility and then can sanction that action. The norm then can be identified from the sanction that has emerged. The would mean that the architectures for norm learning should first include a mechanism for the emergence of sanctions and then apply the mechanisms to learn the norm.

3) Role of dynamic network topology: Norm learning of an agent is impacted by the topology of the connections it has in a networked society. The role of network topology on norm emergence on both static and dynamic network topologies has been studied by many researchers [12], [21], [22], [24], [27], [31]. A particular focus on dynamic network topology will be desirable in the future.

4) Consideration of a noisy environment: Only few research works such as the one by Hoffmann [14] have considered the possibility of a noisy environment. This could be easily incorporated in a learning architecture by including a parameter that governs the noise level in the society.

VI. CONCLUSION

This paper aims at providing an overview of the learning mechanisms used by researchers in the field of normative multi-agent systems. First, it discusses the current state of the field of norm learning. Second, it discusses the three aspects of norm learning namely learning by doing, observing and communicating. It also discusses the need for combining the three aspects of learning and using a simple example it demonstrates how these three aspects can be combined. Third, it compares two architectures for norm learning and also provides pointers to the features that can be incorporated into the norm learning architectures in the future.

ACKNOWLEDGMENT

The author would like to thank Stephen Cranefield, Maryam Purvis and Martin Purvis who may have influenced some of the ideas presented in this paper through the discussions the author has had with them over the years.

REFERENCES

[1] Giulia Andrighetto, Marco Campenni, Federico Cecconi, and Rosaria Conte. The complex loop of norm emergence: A simulation model. In Shu-Heng Chen, Claudio Cioffi-Revilla, Nigel Gilbert, Hajime Kita, Takao Terano, Keiki Takadama, Claudio Cioffi-Revilla, and Guillaume Deffuant, editors, *Simulating Interacting Agents and Social Phenomena*, volume 7 of *Agent-Based Social Systems*, pages 19–35. Springer, 2010.

[2] Giulia Andrighetto, Rosaria Conte, Paolo Turrini, and Mario Paolucci. Emergence in the loop: Simulating the two way dynamics of norm innovation. In Guido Boella, Leon van der Torre, and Harko Verhagen, editors, *Normative Multi-agent Systems*, number 07122 in Dagstuhl Seminar Proceedings. Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany, 2007.

[3] Robert Axelrod. An evolutionary approach to norms. *The American Political Science Review*, 80(4):1095–1111, 1986.

[4] Guido Boella, Leendert van der Torre, and Harko Verhagen. Introduction to the special issue on normative multiagent systems. *Autonomous Agents and Multi-Agent Systems*, 17(1):1–10, 2008.

[5] Magnus Boman. Norms in artificial decision making. *Artificial Intelligence and Law*, 7(1):17–35, 1999.

[6] Marco Campenní, Giulia Andrighetto, Federico Cecconi, and Rosaria Conte. Normal = Normative? The role of intelligent agents in norm innovation. *Mind & Society*, 8(2):153–172, 2009.

[7] Cristiano Castelfranchi, Rosaria Conte, and Mario Paolucci. Normative reputation and the costs of compliance. *Journal of Artificial Societies and Social Simulation*, 1(3), 1998.

[8] Aaron Ceglar and John F. Roddick. Association mining. *ACM Computing Surveys*, 38(2):5, 2006.

[9] Jon Elster. Social norms and economic theory. *The Journal of Economic Perspectives*, 3(4):99–117, 1989.

[10] EMIL project report. Emil - emergence in the loop: simulating the two way dynamics of norm innovation. Emil-t, deliverable 5.1, final project report, EMIL consortium, 2010.

[11] Joshua M. Epstein. Learning to be thoughtless: Social norms and individual computation. *Computational Economics*, 18(1):9–24, 2001.

[12] Nathan Griffiths and Michael Luck. Norm emergence in tag-based cooperation. In *Proceeding of the AAMAS 2010 workshop on Coordination, Organization, Institutions and Norms in agent systems (COIN)*, pages 80–87, 2010.

[13] Mohamed Hamada. Web-based environment for active computing learners. In *ICCSA (1)*, pages 516–529, 2008.

[14] Mathew J. Hoffmann. 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.

[15] Andrew J. I. Jones and Marek Sergot. On the characterisation of law and computer systems: The normative systems perspective. In *Deontic Logic in Computer Science: Normative System Specification*, pages 275–307. John Wiley and Sons, 1993.

[16] L.P. Kaelbling, M.L. Littman, and A.R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1-2):99–134, 1998.

[17] James. E. Kittock. Emergent conventions and the structure of multi-agent systems. In Lynn Nadel and Daniel L. Stein, editors, *1993 Lectures in Complex Systems, Santa Fe Institute Studies in the Sciences of Complexity*. Addison-Wesley, 1995.

[18] Ulf Lotzmann and Michael Möhring. Simulating norm formation: an operational approach. In *Proceedings of the Eighth International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 1323–1324. International Foundation for Autonomous Agents and Multiagent Systems, 2009.

[19] S. Muggleton and L. De Raedt. Inductive logic programming: Theory and methods. *The Journal of Logic Programming*, 19:629–679, 1994.

[20] Martin Neumann. A classification of normative architectures. In Keiki Takadama, Claudio Cioffi-Revilla, and Guillaume Deffuant, editors, *Simulating Interacting Agents and Social Phenomena: The Second World Congress*, volume 7 of *Agent-Based Social Systems*, pages 3–18. Springer, Berlin/Heidelberg, 2010.

[21] Josep M. Pujol. *Structure in Artificial Societies*. PhD thesis, Departament de Llenguatges i Sistemes Informàtics, Universitat Politènica de Catalunya, 2006.

[22] Norman Salazar, Juan A. Rodríguez-Aguilar, and Josep L. Arcos. Infection-based norm emergence in multi-agent complex networks. In *Proceeding of the 2008 conference on ECAI 2008*, pages 891–892, Amsterdam, The Netherlands, The Netherlands, 2008. IOS Press.

[23] Bastin Tony Roy Savarimuthu and Stephen Cranefield. A categorization of simulation works on norms. In Guido Boella, Pablo Noriega, Gabriella Pigozzi, and Harko Verhagen, editors, *Normative Multi-Agent Systems*, number 09121 in Dagstuhl Seminar Proceedings, pages 43–62, Dagstuhl, Germany, 2009. Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany.

[24] Bastin Tony Roy Savarimuthu, Stephen Cranefield, Maryam Purvis, and Martin K. Purvis. Role model based mechanism for norm emergence in artificial agent societies. In *Coordination, Organizations, Institutions, and Norms in Agent Systems III*, pages 203–217, 2007.

[25] Bastin Tony Roy Savarimuthu, Stephen Cranefield, Maryam A. Purvis, and Martin K. Purvis. Norm identification in multi-agent societies. Discussion Paper 2010/03, Department of Information Science, University of Otago, 2010.

[26] Bastin Tony Roy Savarimuthu, Stephen Cranefield, Maryam A. Purvis, and Martin K. Purvis. Obligation norm identification in agent societies. *Journal of Artificial Societies and Social Simulation*, 13(4):3, 2010.

[27] Sandip Sen and Stephane Airiau. Emergence of norms through social learning. In *Proceedings of the Twentieth International Joint Conference on Artificial Intelligence (IJCAI)*, pages 1507–1512. AAAI Press, 2007.

[28] Yoav Shoham and Moshe Tennenholtz. Emergent conventions in multi-agent systems: Initial experimental results and observations. In *Proceedings of the Third International Conference on the Principles of Knowledge Representation and Reasoning (KR)*, pages 225–231, San Mateo, CA, USA, 1992. Morgan Kaufmann.

[29] Stanislaw A. B. Stane and Mariusz Zytniewsk. Normative multi-agent enriched data mining to support e-citizens. In Longbing Cao, editor, *Data Mining and Multi-agent Integration*, pages 291–304. Springer, 2009.

[30] Harko Verhagen. Simulation of the Learning of Norms. *Social Science Computer Review*, 19(3):296–306, 2001.

[31] Daniel Villatoro, Sandip Sen, and Jordi Sabater-Mir. Topology and memory effect on convention emergence. In *WI-IAT '09: Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, pages 233–240, Washington, DC, USA, 2009. IEEE Computer Society.

[32] Adam Walker and Michael Wooldridge. Understanding the emergence of conventions in multi-agent systems. In Victor Lesser, editor, *Proceedings of the First International Conference on Multi-Agent Systems (ICMAS)*, pages 384–389, Menlo Park, California, USA, 1995. AAAI Press.

[33] Christopher J. C. H. Watkins and Peter Dayan. Q-learning. *Machine Learning*, 8(3-4):279–292, 1992.

[34] Roel J. Wieringa and John-Jules Ch. Meyer. Applications of deontic logic in computer science: a concise overview. In *Deontic logic in computer science: Normative system specification*, pages 17–40. John Wiley & Sons, Inc., New York, USA, 1994.