Analyzing the tradeoff between efficiency and cost of norm enforcement in stochastic environments populated with self-interested agents

Moser Silva Fagundes[1] and **Sascha Ossowski**[2] and **Felipe Meneguzzi**[3]

Abstract. In multiagent systems, agents might interfere with each other as a side-effect of their activities. One approach to coordinating these agents is to restrict their activities by means of social norms whose violation results in sanctions to violating agents. We formalize a normative system within a stochastic environment and norm enforcement follows a stochastic model in which stricter enforcement entails higher cost. Within this type of system, we provide an approach to analize the tradeoff between norm enforcement efficiency and its cost considering a population of norm-aware selfish agents.

1 INTRODUCTION

In this paper, we define the dynamics of the world and the norm enforcement mechanism without the assumption of determinism, and we formalize the decision making process of our autonomous agents with *Normative Markov Decision Processes* (NMDPs) [1, 2]. Such a stochastic norm-based coordination mechanism aims at ensuring that the multiagent system as a whole runs according to the properties specified in a set of social norms. Thus, agents are subjected to social norms that dictate prohibitions and obligations with regard to world-states, and failure to comply with these social norms brings about some kind of sanction. Norms are enforced on the basis of the observation of the current world-state by a stochastic mechanism that detects violations with a certain probability. Within this model, we associate a cost to norm enforcement such that stricter enforcement mechanisms incur a higher cost to the enforcement authority. Such an enforcement costing model mirrors the real world, where mechanisms with a higher probability of detecting (and sanctioning) violations are more expensive. We develop a simulation-based method to calculate this tradeoff, and show results for a representative scenario. In summary, this paper makes two major contributions. First, we define a rich stochastic norm enforcement mechanism in stochastic environments populated with selfish rational agents, and second, we provide insights into the tradeoffs involved in norm enforcement.

2 NORMATIVE STOCHASTIC ENVIRONMENT

Let $\mathcal{G} = \{a_1, \ldots, a_n\}$ be the set of agents, modeled as NMDPs [1, 2], operating in the multiagent system. The state space of an agent $a_i \in \mathcal{G}$ is defined as follows. Let $\mathcal{F}$ be the set of features that characterize different aspects of the states of the world. A feature $f_i \in \mathcal{F}$ can take on a finite number of values and $\mathcal{V}_{f_i}$ corresponds to the finite set of possible values of f_i. The state of an agent is a complete assignment of values to its features, and the state space $\mathcal{S}$ is the cross product of the value spaces for the features, i.e.: $\mathcal{S} = \times_{i=1}^{|\mathcal{F}|} \mathcal{V}_{f_i}$. The *current state* of the agents determines the system's current state and it is represented as a vector $\{s_1, \ldots, s_n\}$ in which s_i is the current state of $a_i \in \mathcal{G}$.

Definition 1 (Norm) *A norm is a tuple* $\langle \delta, \mathcal{X}, \mathcal{E}, \sigma \rangle$ *where the following constraints hold:* $\delta \in \{obligation, prohibition\}$ *is the deontic modality;* $\mathcal{X} \subseteq \mathcal{S}$ *is the set of states (normative context) in which the norm applies;* $\mathcal{E} \subseteq \mathcal{X}$ *is the subset of states in the normative context which are obliged or prohibited (target states); and* σ *is a sanction represented by a tuple* $\langle \rho, \phi \rangle$ *where:*

- $\rho : \mathcal{S} \mapsto \mathbb{R}$ *is a function that gives the penalty for violating this norm in a given state* ($\rho(s)$ *yields the penalty to be paid in* s);
- $\phi : \mathcal{S} \mapsto \mathcal{S}$ *is a function that calculates the state resulting from an enforced state-transition in response to the violation of this norm* ($\phi(s)$ *yields the outcome of an enforced transition in* s).

A set of states is relevant to a norm $\langle \delta, \mathcal{X}, \mathcal{E}, \sigma \rangle$ if this set of states is a subset of $\mathcal{X}$, which indicates the context where the norm applies. Given a set of states that are relevant to a norm, we can determine which of them violate it.

Definition 2 (Violating states) *Let* $q \in \mathcal{N}$ *be a norm* $\langle \delta, \mathcal{X}, \mathcal{E}, \sigma \rangle$, *the set of states that violate* q, *denoted as* $\mathcal{S}_q^{\triangledown}$, *is defined as:*

$$\mathcal{S}_q^{\triangledown} = \begin{cases} \mathcal{E} & if \ (\delta = prohibition) \\ \mathcal{X} \setminus \mathcal{E} & if \ (\delta = obligation). \end{cases}$$

The model of norm enforcement mechanism developed in this paper is based on the detection of violating states in terms of observations of the agents' current state in the world. Such observations are assumed to be imperfect, so that the mechanism only detects violations with a certain probability, as stated in Definition 3.

Definition 3 (Detection model) *Let* $\mathcal{N}$ *be the set of norms and* $\mathcal{S}$ *be the state space. A probabilistic detection model consists of a function* $\mathcal{D} : \mathcal{N} \times \mathcal{S} \rightarrow [0,1]$ *such that* $\mathcal{D}(q, s)$ *returns the detection probability of the violation of the norm* $q \in \mathcal{N}$ *in the state* $s \in \mathcal{S}$.

Besides being imperfect, the enforcement mechanism is *resource-bounded* so that monitoring the environment has an associated cost.

[1] Federal Institute of Education, Science and Technology Sul-Rio-Grandense (IFSul), RS, Brazil, email: moserfagundes@charqueadas.ifsul.edu.br
[2] Centre for Intelligent Information Technologies, University Rey Juan Carlos (URJC), Móstoles, Madrid, Spain, email: sascha.ossowski@urjc.es
[3] School of Informatics, Pontifical Catholic University of Rio Grande do Sul (PUCRS), Porto Alegre, Brazil, email: felipe.meneguzzi@pucrs.br

This cost is a function of the accuracy of observations and the size of the population of agents. We formalize this function as $\mathcal{MK}(\mathcal{D}, \mu)$, which returns the cost per time step of detecting violations according to the model $\mathcal{D}$ in a multiagent system with population size μ.

Similarly to detecting norm violations, the execution of sanctions implies costs for the norm enforcement mechanism. Thus, the sanctioning cost $\mathcal{SK}(t)$ for a given time step t, is:

$$\mathcal{SK}(t) = \sum_{i=1}^{|\mathcal{N}|} \mathcal{I}(\sigma_i, t)\ \mathcal{SSK}(\sigma_i)$$

where $\mathcal{I}(\sigma_i, t)$ gives the number of times that the sanction σ_i has been imposed in the time step t, and $\mathcal{SSK}(\sigma_i)$ returns the cost of executing the sanction σ_i. In summary, the total enforcement cost at a given time step t, denoted as $\mathcal{K}(t)$, is determined as follows:

$$\mathcal{K}(t) = \mathcal{MK}(\mathcal{D}, \mu) + \mathcal{SK}(t). \tag{1}$$

3 EXPERIMENTS

Our approach aims to determine effective enforcement intensities that balance the ability of the mechanism to self-support and at the same time to ensure that coordination problems do not exceed a maximum acceptable level. To illutrate our approach, we have used a motion environment made up of lanes composed of discrete contiguous cells, where the agents are able to move one cell at a time. There are gateways through which the agents enter and leave the environment. Non-determinism is modeled by the fact the actions are "unreliable" – their intended outcome occurs with probability 0.99, but with probability 0.01 the agent remains in the same state.

In our experiments, the agents know the norms and the detection probabilities. That is, although they do not know whether individual norm violations will be detected, they are aware of the probability of violations being detected in general. The agents' perception is incomplete, so while traveling across the environment, the agents always know their own position, but not the position of other agents. This assumption may cause coordination problems. To cope with these problems, we introduce a set of norms to regulate the traveling direction in each lane. If a violation is detected, the agent is driven to the obliged direction and loses $0.1\mathsf{R_N}$, where $\mathsf{R_N}$ is the reward unit. These norms are specified using the following template, where X is a particular lane and Y is the obliged moving direction:

$$\langle \text{OBLIGATION}, (\text{LANE=X}), (\text{DIRECTION=Y}),\\ \langle \{ \top \rightarrow -1.0\mathsf{R_N} \},\{ \top \rightarrow \{(\text{DIRECTION=Y})\} \} \rangle\rangle$$

Origin and destination gateways of each agent were randomly assigned using a uniform distribution. Each simulation ran for 10^6 time steps, allowing us to identify the parameters and value ranges in which the societal behavior changes significantly:

- $\mathcal{D}(q, s) = \beta$ is the detection probability of norm violations for all norms and all states, where β ranges from 0.01 to 0.13.
- μ is the population setting, which is the number of agents in the environment at any given time throughout the entire simulation.

The norm enforcement cost is computed using Formula 1 where $\mathcal{MK}(\mathcal{D}, \mu)$ is the cost of monitoring norm violations, and $\mathcal{SK}(t)$ is the sanctioning cost. The monitoring cost and the sanctioning cost in our motion environment are defined, respectively, as follows:

$$\mathcal{MK}(\mathcal{D}, \mu) = \mathcal{MK}(\beta, \mu) = 1.1^{10\mu\beta} \times 10^{-3}\mathsf{R_N}$$

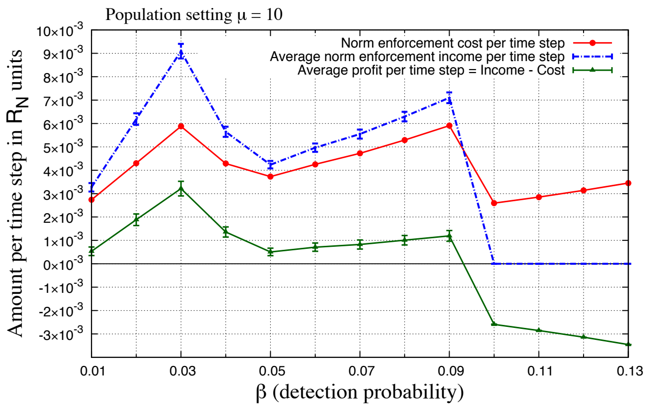

Figure 1. Average norm enforcement cost, income and profit per time step.

$$\mathcal{SK}(t) = \sum_{i=1}^{8} \mathcal{I}(\sigma_i, t)\ 0.05\mathsf{R_N}$$

where $0.05\mathsf{R_N}$ is the cost of executing any sanction (there are 8 obligations, one per lane, and the cost of sanctioning in each of them is the same). Note that each cost function is domain-specific. For example, in this work, the monitoring cost increases exponentially with β and μ, and the cost of executing a sanction is fixed.

The norm enforcement cost per time step $\mathcal{K}(t)$ is drawn in Figure 1 as a solid line with circles. If we subtract this cost from the average enforcement income (dotted line), we have the average profit (solid line with triangles). In our experiments, the system profits when $\beta \lessapprox 0.09$. Notice that the enforcement income increases as μ (number of agents simultaneously running in the environment) increases. Assuming that the detection cost does not depend on the value of μ, certain normative solutions may become profitable if we allow more agents to join the system. Although Figure 1 shows that the highest profit occurs when $\beta = 0.03$, in systems with self-interested agents, such a low detection probability of violations may entail a significant amount of coordination problems, nullifying the effectiveness of the norms as a multiagent coordination device. Based on the information provided by the simulation, the designer of the norms is able to analyze the tradeoff between efficiency and cost of norm enforcement.

ACKNOWLEDGEMENTS

This research is partially supported by the Spanish Ministry of Science and Innovation through the project AT (grant CSD2007-0022; CONSOLIDER-INGENIO 2010) and by the Spanish Ministry of Economy and Competitiveness through the project iHAS (TIN2012-36586-C03-02). The second author thanks FAPERGS for the financial support within process 12/0808-5 under the ARD program.

REFERENCES

[1] Moser Silva Fagundes, Holger Billhardt, and Sascha Ossowski, 'Reasoning about Norm Compliance with Rational Agents', in *ECAI*, volume 215 of *Frontiers in Artificial Intelligence and Applications*, pp. 1027–1028. IOS Press, (2010).

[2] Moser Silva Fagundes, Sascha Ossowski, Michael Luck, and Simon Miles, 'Using Normative Markov Decision Processes for evaluating electronic contracts', *AI Commun.*, **25**(1), 1–17, (2012).