Runtime Revision of Norms and Sanctions based on Agent Preferences (Extended Abstract)*,**

Davide Dell'Anna, Mehdi Dastani, and Fabiano Dalpiaz

Utrecht University, Utrecht, The Netherlands
{d.dellanna, m.m.dastani, f.dalpiaz}@uu.nl

CEUR-WS.org/Vol-2491/abstract43.pdf

A multiagent system (MAS) comprises a set of autonomous agents that interact in a shared environment [9]. For example, a smart traffic system is a MAS that includes autonomous agents like cars, traffic lights, etc. The objectives of such a system include ensuring that each agent reaches its destination timely, and that the number of accidents is minimized.

For a MAS to achieve its system-level objectives, the complexity and unpredictability of the agent interactions and of the environment must be taken into account. When engineering such systems, the available knowledge of these dynamics is only partial and incomplete. Therefore, MASs need to be regulated at runtime. Norm enforcement is a prominent mechanism for controlling and coordinating the runtime behavior of the agents in a MAS without over-constraining their autonomy [1,8]. Norm enforcement via sanctions is traditionally contrasted with norm regimentation, which prevents the agents from reaching certain states of affairs. For example, in a smart traffic system, a regimentation strategy could be to close a road to prevent cars from entering the road, while a sanctioning strategy could be to impose sanctions on cars that do enter the road.

Existing research has studied the offline revision of the enforced norms, and proposed logics that support norm change [2, 7], and studied the legal effects of norm change [6]. In previous work [4] we have proposed a framework for engineering normative MASs that, using runtime data from MAS execution, revises the norms in the MAS to maximize the achievement of the system objectives. In such work, we made the simplistic assumption that norms are regimented and we introduced algorithms for switching among alternative norms.

In this paper, we make a step forward toward the engineering of normative multiagent systems and we propose a regulatory mechanism that relies on *norms with sanction*. In our approach, we automatically revise the sanctions that are employed to enforce the norms. To do so, we first interpret—through a Bayesian Network—runtime execution data in terms of how well certain norms contribute to the achievement of the system-level objectives in different operating contexts. Then, we suggest a revision of the sanctions using two different heuristic strategies, called SYNERGY and SENSITIVITY. The two strategies leverage the knowledge learned from runtime execution data and the knowledge about the

* Copyright © 2019 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

** The full version of the paper appears in [5].

preferences of the agents in the MAS. We assume preferences to be specified in terms of a desired state of affairs and the maximum payment that the agent is willing to make to achieve the state of affairs. Agents' preferences provide an estimation of the upper bound of the probability of violating a norm, given a certain population of agents. Runtime execution data provides information about the exhibited behavior of the agents (i.e., the exhibited probability of violating a norm). Synergy and Sensitivity compare these two pieces of information and determine new sanctions to use to enforce a norm in different operating contexts.

Synergy: the new sanction is the closest sanction to the current one that is expected to decrease the probability of violating a norm, in case of positive synergy between norm and system-objectives (i.e., objectives are more likely achieved when the norm is obeyed), or to increase it, in case of negative synergy.

Sensitivity not only determines the direction of the revision–i.e., increasing or decreasing the probability of violating the norm–, but also provides an estimation of the *required change* in such probability in order to make the norm effective for the achievement of the system-objectives. Such estimation is obtained by performing sensitivity analysis for the Bayesian Network [3]. The new sanction is then the sanction that is expected to change the probability of violating a norm as required by Sensitivity.

An evaluation through a traffic regulation simulation shows that our heuristics outperform uninformed heuristics in terms of how fast they identify an optimal solution (i.e., an optimal sanction for the enforced norm). We perform six experiments, differing in the distribution of types of agents and in the norm enforced in the MAS. We show that in all experiments our informed heuristics, and especially Sensitivity, provide a significant improvement in the number of steps required to identify optimal sanctions (reducing the required number of steps up to 98%), compared to the uninformed strategies.

References

1. Alechina, N., Bulling, N., Dastani, M., Logan, B.: Practical run-time norm enforcement with bounded lookahead. In: Proc. of AAMAS. pp. 443–451 (2015)
2. Aucher, G., Grossi, D., Herzig, A., Lorini, E.: Dynamic context logic. In: Proc. of LORI. pp. 15–26 (2009)
3. Chan, H., Darwiche, A.: Sensitivity analysis in bayesian networks: From single to multiple parameters. In: Proc. of UAI. pp. 67–75 (2004)
4. Dell'Anna, D., Dastani, M., Dalpiaz, F.: Runtime norm revision using bayesian networks. In: Proc. of PRIMA. pp. 279–295 (2018)
5. Dell'Anna, D., Dastani, M., Dalpiaz, F.: Runtime revision of norms and sanctions based on agent preferences. In: Proc. of AAMAS. pp. 1609–1617 (2019)
6. Governatori, G., Rotolo, A.: Changing legal systems: legal abrogations and annulments in defeasible logic. Logic Journal of the IGPL **18**(1), 157–194 (2010)
7. Knobbout, M., Dastani, M., Meyer, J.C.: A dynamic logic of norm change. In: Proc. of ECAI. pp. 886–894 (2016)
8. Testerink, B., Dastani, M., Bulling, N.: Distributed controllers for norm enforcement. In: Proc. of ECAI. pp. 751–759 (2016)
9. Wooldridge, M.J.: An Introduction to MultiAgent Systems (2. ed.). Wiley (2009)