Reasoning with Agent Preferences in Normative Multi-agent Systems

(Extended Abstract)

Jie Jiang
TU Delft
Delft, the Netherlands
jie.jiang@tudelft.nl

John Thangarajah
RMIT
Melbourne, Australia
john.thangarajah@rmit.edu.au

Huib Aldewereld
TU Delft
Delft, the Netherlands
h.m.aldewereld@tudelft.nl

Virginia Dignum
TU Delft
Delft, the Netherlands
m.v.dignum@tudelft.nl

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Languages and structures*; I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Multiagent systems*

General Terms

Design, Verification

Keywords

Normative systems; Agent preferences

1. INTRODUCTION

A fundamental feature of autonomous agents is the ability to make decisions in dynamic environments where alternative plans may be used to achieve their goals. From an individual perspective, agents can have personal preferences over some actions and they try to maximize their preference satisfaction in deciding the actions to achieve their goals. In normative multi-agent systems, however, agent actions are not only directed by their own personal preferences but also the normative constraints imposed by the system. Within this context, the agents decide on their actions based on the reasoning of (1) whether their actions violate the norms imposed by the system, and (2) to what extent their actions satisfy their individual preferences. As such, this necessitates mechanisms to provide the agents with information about both the normative consequence and the preference satisfaction of their actions in an integrated way.

In this paper, we propose a unified framework to analyze agent interactions taking into consideration the agent preferences in the setting of normative multi-agent systems. To reason about the normative consequence, we use the normative language presented in [1]. To reason with agent preferences, we extend the work of Visser et al. [2] such that it allows us to specify the agents' preferences over their own actions, other agents' actions, and actions relating to roles. We further extend the computational model presented in [1] such that we can obtain a quantitative measurement of the normative consequence and the preference satisfaction with respect to the agents' actions in a unified way.

This approach can be used for individual agents to reason about their actions considering their personal preferences as well as the normative constraints. If agents are willing to disclose their preference information, the approach can also be used to evaluate agent interactions from a system point of view, such as maximizing the average satisfaction of the whole population of the agents.

Appears in: *Alessio Lomuscio, Paul Scerri, Ana Bazzan, and Michael Huhns (eds.), Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014), May 5-9, 2014, Paris, France.*
Copyright © 2014, International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

2. THE FRAMEWORK

Within the context of a normative multi-agent system, our framework consists of two modules: specification and evaluation. The *specification module* specifies the normative constraints imposed by the system, the individual preferences of the agents in the system, and the possible interaction plans of the agents. Given the specification module, the *evaluation module* will evaluate each plan following two parallel steps: (1) the *compliance evaluation* is to verify the plan against the normative constraints and determine the compliance status of the plan, and (2) the *satisfaction evaluation* is to verify the plan against the preferences of each agent and indicate to what extent the agent is satisfied with the plan.

Finally, we obtain an integrated picture of the compliance status and the satisfaction level of all the participating agents with respect to each plan. From an individual perspective, this equips the agents with adequate information to reason about their actions on the basis of both the normative consequence and the personal satisfaction. From a system point of view, this provides the possibility to identify plans that are favored by the individual agents provided that the plans are in accordance with the system constraints.

2.1 Specification Module

Normative constraints: The specification of normative constraints is realized by using the normative language proposed in [1]. It provides the components to specify norms (obligations and prohibitions) as well as their compliance relations (e.g., choice and reparation).

Agent preferences: The specification of agent preferences is realized by extending the preference language proposed in [2] with the following adjustments. Firstly, agent preferences are specified over actions. For example, *Alice* prefers to book flights herself (*bookFS*). Secondly, the agents' preferences are not restricted to their own actions but might depend on other agents' actions. For example, *Bob* has a preference of booking flights by the travel agency (*bookFA*) over booking train tickets (*bookT*) if the other two group members *Alice* and *Carl* choose to *bookFA*. Thirdly, an agent's preference might depend on the actions of some other agents enacting a particular role in the system. For example, *Carl* might prefer *bookFS* over *bookT* if the group *leader* choose to *bookFS*. Similarly, we use *basic desire formulas* to represent basic statements about the preferred situation, *atomic preference formulas* to represent an ordering over basic desire formulas and *general preference formulas* to express atomic preference formulas that are optionally subjected to a condition. The preferences of an agent are specified as a set of general preference formulas which in our approach are built over agent actions and role actions using numeric values to indicate different levels of preferences.

Interaction plans: In our setting, agent interactions are captured by sequences of agent actions, denoted as interaction plans (IP). Figure 1 shows an example of three possible IPs involving five agents (an organization X, three employees *Alice*, *Bob*, *Carl*, and a software system *Sys*). The three employees are together granted a group trip by X and they have to make requests (*request*), book hotels (*book3*$\star$, *book4*$\star$, *book5*$\star$), make declarations (*declare*), while *Sys* will inform declarations (*inform_D*), inform extra cost (*inform_E*), inform violations (*inform_V*). The ordering of the agent actions is indicated by their positions relative to the time line at the bottom.

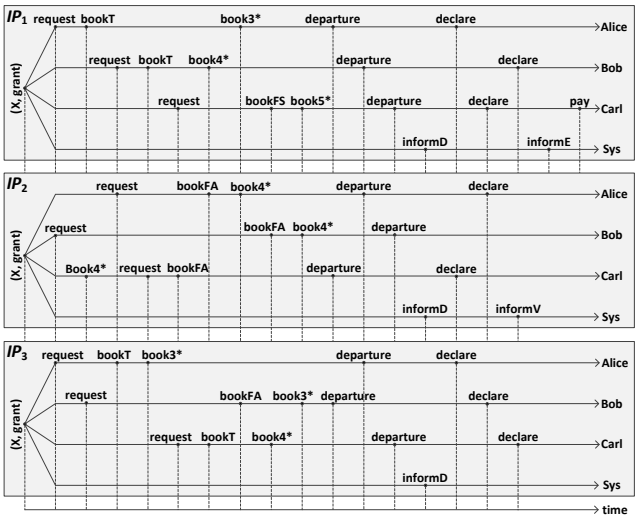

Figure 1: Interaction Plans

2.2 Evaluation Module

Based on the specification module, the evaluation module will take two evaluation steps and provide the information on (1) compliance status, i.e., whether the plan violates the norms imposed by the system, and (2) preference satisfaction, i.e., to what extent the plan satisfies the preferences of the individual agents. An extension to the computational model for normative evaluation in [1] allows for the calculation of them both. Compliance status is calculated for each plan and is expressed by labeling each plan as full compliance (FC), sub-ideal compliance (SC) or non compliance (NC). Preference satisfaction is calculated for each agent with respect to each plan and is expressed by a numeric value.

Table 1 shows the evaluation results of the three plans described in Figure 1, according to the organization's normative constraints and the agents' individual preferences (which are left out due to space limitations). It can be seen that the third plan is mostly favorable in terms of the normative consequence. According to the satisfaction level, *Alice* prefers the second plan, *Bob* prefers the first plan, and *Carl* prefers the second plan.

interaction plans	compliance status	Preference satisfaction		
		Alice	Bob	Carl
IP_1	SC	30	0	50
IP_2	NC	25	100	10
IP_3	FC	30	18	170

Table 1: Evaluation of Interaction Plans

Whether a plan is indeed the choice of an agent depends on the combined-reasoning strategy of the agent. For example, if *Alice* is *selfish*, her choice would be the second plan since she will try to maximize her preference satisfaction without considering the normative consequence. If *Alice* is *norm-aware*, the third plan would be her choice since she will try to minimize the violations. In regulated systems, norm compliance is an important feature for interaction plans since violations may cause a failure to the system as a whole. Therefore, from a system point of view, a possible combined-reasoning strategy is to maximize the average satisfaction of the whole group and at the same time to ensure norm compliance. In this case, the first plan would be a good choice for the whole group.

3. CONCLUSIONS

In this paper, we explored the integration of agent preferences with normative multi-agent systems such that agents can reason about their actions with the information of both preference satisfaction and norm compliance in a unified framework. For future work, we seek to enrich our preference language, e.g., specifying preferences by considering the resource usage of actions. Moreover, we intend to further investigate the interplay between norm compliance and agent preferences, and build formalisms to enable preference specifications over normative consequences.

4. REFERENCES

[1] J. Jiang, H. Aldewereld, V. Dignum, and Y.-H. Tan. Norm contextualization. In *Proceedings of Coordination, Organizations, Institutions, and Norms in Agent Systems VIII*, LNCS, pages 141–157, 2013.

[2] S. Visser, J. Thangarajah, and J. Harland. Reasoning about preferences in intelligent agent systems. In *Proceedings of International Conference on Artificial Intelligence*, pages 426–431, 2011.