Normative Conflict Detection and Resolution in Cooperating Institutions

Tingting Li *
Dept. of Computer Science
University of Bath, United Kingdom
t.li@cs.bath.ac.uk

Abstract

Institutions (also called normative frameworks) provide an effective mechanism to govern agents in open distributed systems. An institution specifies a set of norms, with respect to the achievement of a goal or goals, that regulate agents' behaviours in terms of permissions, empowerments and obligations. However, in most real circumstances, several institutions probably have to cooperate to govern the same entities simultaneously, which is very likely to give rise to norm conflicts simply if institutions will be designed independently and typically with different goals. In this thesis, we aim: (i) to identify the different ways to combine institutions, (ii) to model those ways formally and computationally by extending an existing model for single institutions, (iii) to detect conflicts in different types of combined institutions automatically, and (iv) to resolve those conflicts via automatic norm revision using an approach based on inductive learning.

1 Introduction

Regulating agents' behaviours in an open environments is often a troublesome problem because agents are normally self-interested and heterogeneous. However, agents' behaviours can be constrained and regulated for the achievedment of certain objectives, by means of institutions, in which a set of norms is specified according to the objectives. The norms often specify under which kind of circumstances, which actions are permitted, prohibited or obliged to perform for an agent. Institutions have long been investigated in the normative agent community but research so far has mainly concentrated on either single institutions or multiple homogeneous institutions which are designed for a particular system and purpose (examples are [Cliffe *et al.*, 2007; Noriega, 1997; Vázquez-Salceda, 2003]). However, in most real circumstances, it is very likely that institutions that are designed independently have to cooperate together to govern actors' behaviours. Under these circumstances, conflicting norms are very likely to occur because the institutions are established by different designers with distinct objectives and purposes. When an event is prohibited (not permitted) by one institution, but simultaneously permitted or obligated by another, we identify this situation as a norm conflict in a cooperating institution. Different structures of cooperating institutions may suggest different types and natures of conflicts and hence require different strategies for detecting and resolving conflicts. In the thesis, we firstly model three different ways of forming cooperating institutions formally and computationally. Based on those models, we discuss our mechanism for automatic conflict detection and resolution in detail.

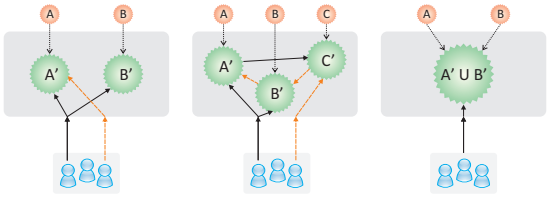

Figure 1: Three cooperating institution types: a) composite, b) multi-, and c) merged.

For context, we first briefly consider how heterogenous institutions may be combined to form cooperating institutions, as shown in Fig. 1, namely: (a) the *composite institution*: the institutions are treated individually, which implies that a composite institution is not a new institution, but rather a common governance scope of all the individual institutions. The key problem of this structure is how to make such a combination consistent while maintaining the autonomy of each individual institution. (b) the *multi-institution*: introduced by [Cliffe *et al.*, 2007], where a hierarchical structure of interlinked institutions is suggested, in which, an institution might be governed by other institutions, and thereby the states of the inferior institution might be updated by not only the external events, but also by the superior institutions. (c) the *merged institution*: in which all the norms of each institutions are merged to form a completely new conflict-free "super-institution". In Fig.1, the A', B', C' and $A' \cup B'$ denote the consistent states of the institutions or their consistent merged union. The arrows show triggered events and changes to the state as a consequence of the actions performed by two agents, differentiated by solid and dashed lines.

*Supervised by: Julian Padget and Marina De Vos, University of Bath, UK. Special thanks to Tina Balke, University of Surrey, UK and Ken Satoh, NII, Japan.

In order to automate conflict detection and resolution, we need to model institutions formally and computationally, which then enable us to analyze the evolution of institutional states (i.e.normative information) and thereby examine potential conflicts. We employ the event-driven formalism Inst*AL* presented by [Cliffe *et al.*, 2007], which is a formal declarative language that models the change of institutional states as a result of the interpretation of external events. The formal model is then converted into a logic program under Answer Set semantics [Gelfond and Lifschitz, 1991]. We extend the single institutional model to express the three cooperating institution structures identified above. The transition of institutional states are driven by event traces, and thus the conflicts can be detected in the states of any pair of institutions where one fluent (i.e. the permission, empowerment or obligation towards a particular event) is true in one and false in the other. Having detected the pairwise conflicts, the thesis then proposes an approach based on *Inductive Logic Programming* [Corapi *et al.*, 2011] to produce appropriate revisions to the norms of the less important institution of the conflict pair, in order to be consistent with the others.

2 Progress to Date

So far we have been focusing on the first type of cooperating institution: composite institutions. We propose a conflict detection mechanism [Li *et al.*, 2012a; 2012b] which is able to: (i) in general identify all the potential conflicts among institutions of a cooperating institution by examining all possible event traces in the context of the cooperating institution, and also (ii) detect conflicts caused by one particular course of events, which is of particular interests for users. Together with the detection mechanism, we extend our definition of conflicts to differentiate *weak conflicts* – between permission and prohibition (not permitted) – and *strong conflicts* – between obligation and prohibition.

Regarding conflict resolution, the thesis aims to revise norms of one or more institutions in respect of the conflicts detected between them. We treat the problem as a *Theory Revision* task which can be solved by inductive learning approaches [Corapi *et al.*, 2011]. Our mechanism supports not only the synthesis of new norms but also deletion or revision of existing norms by means of use cases. A use case is a sequence of events that lead to conflicts among the combined institutions. With a set of conflicting institutions, the one with the least precedence will need to be revised, whose norms thus are transformed as an hypothesis to be searched for revision, while the other higher-precedence institutions will be part of the background theory of the learning task. The whole methodology is not only presented in theory, but also implemented automatically by *Answer Set Programming* to be applicable with the computational models of institutions. The final produced solutions could be atomic, partial or complete solutions, which in order suggests revisions to the norms such as to resolve one, several or all conflicts towards a particular event trace. To evaluate our approach, we also provide a measurement of the revisions in terms of *number of revision operations*. Thereby, the approach also guarantees to provide the optimal solution as measured by the minimum number of revision operations so that we can keep the revised institutions as similar as possible to the original to minimize the side-effects of the revisions.

We also demonstrate our approach with a case study from the juris-informatics domain [Li *et al.*, 2012b] in which we successfully applied our methodology to detecting and resolving conflicts between laws from different countries, and also discussed the possibility of applying it to detecting conflicts between newly-enacted laws and other existing ones.

3 Future Plans

We plan several lines for our further development from the current work for inclusion in the thesis: (i) we will investigate conflicts in the other two structures for cooperating institutions – multi-institutions and merged institutions, as depicted in Fig.1. (ii) we will also explore further kinds of conflicts, for example, when an event is permitted by one institution but not empowered by another. In addition, the multi-institution structure reveals more interesting types of conflicts. For example, when one inferior institution is governed by two superior institutions, conflicts may occur between the two superior institutions when the change of the inferior institution leads to inconsistent changes to their states.(iii) the current inductive learning algorithm might produce several answer sets, each of which provides a solution to the task. However, not all of them are sensible revisions, and thus mechanisms for further refining the solutions is needed. (iv) the precedence order is currently established over whole institutions (collections of norms). We believe it may be useful to explore a finer-grained approach using precedence ordering over individual norms rather than whole institutions.

References

[Cliffe *et al.*, 2007] O. Cliffe, M. De Vos, and J. Padget. Answer set programming for representing and reasoning about virtual institutions. In K. Inoue, K. Satoh, and F. Toni, editors, *Computational Logic in Multi-Agent Systems*, volume 4371 of *LNCS*, pages 60–79. Springer, 2007.

[Corapi *et al.*, 2011] D. Corapi, A. Russo, M. De Vos, J. Padget, and K. Satoh. Normative design using inductive learning. *TPLP*, 11(4-5):783–799, 2011.

[Gelfond and Lifschitz, 1991] M. Gelfond and V. Lifschitz. Classical negation in logic programs and disjunctive databases. *New Generation Computing*, 9(3-4):365–386, 1991.

[Li *et al.*, 2012a] T. Li, T. Balke, M. De Vos, K. Satoh, and J. Padget. Conflict detection in composite institutions. In *International Workshop on Agent-based Modeling for Policy Engineering (AMPLE 2012)*, page 66, 2012.

[Li *et al.*, 2012b] T. Li, T. Balke, M. De Vos, K. Satoh, and J. Padget. Detecting conflicts in legal systems. In *Proceedings of the Sixth International Workshop on Juris-informatics (JURISIN 2012)*, 2012.

[Noriega, 1997] P. Noriega. *Agent mediated auctions: The Fishmarket Metaphor*. PhD thesis, Universitat Autònoma de Barcelona, 1997.

[Vázquez-Salceda, 2003] J. Vázquez-Salceda. *The role of norms and electronic institutions in multi-agent systems applied to complex domains. The HARMONIA framework*. PhD thesis, Technical University of Catalonia, 2003.