Chapter XV
Verifying Organizations Regulated by Institutions

Francesco Viganò
Università della Svizzera italiana, Switzerland

Marco Colombetti
Università della Svizzera italiana, Switzerland
Politecnico di Milano, Italy

ABSTRACT

Institutions have been proposed to explicitly represent norms in open multi-agent systems, where agents may not follow them and which therefore require mechanisms to detect violations. In doing so, they increase the efficiency of electronic transactions carried out by agents, but raise the problem of ensuring that such institutions are not characterized by contradictory norms, and provide agents with all the needed powers to fulfill their objectives. In this chapter we present a framework to verify organizations regulated by institutions, which is characterized by a precise formalization of institutional concepts, a language to describe institutions, and a tool to model check them. Finally, to evaluate and exemplify our approach, we model and verify the Chaired Meeting Institution, showing that the verification of institutional rules constitutes a necessary step to define sound institutions.

INTRODUCTION

To automate tasks whose performance is regulated by norms, like the allocation of human organs (Vázquez-Salceda, Dignum, & Dignum, 2005) or the bargain of goods in competitive markets (Noriega, 1997), electronic institutions have been proposed as a mechanism to represent norms in agent organizations (Noriega, 1997) where agents are developed by different entities and their internal states are not accessible. The term “electronic institution” is often used to refer to either the rules that regulate open multiagent systems, the organization that enforces them (Esteva, Rodríguez-Aguilar, Sierra, García,
Verifying Organizations Regulated by Institutions

& Arcos, 2001), the software implementation of institutional rules (Esteva, Rodriguez-Aguilar, Rosell, & Arcos, 2004), or a specific formalism to describe them (Esteva, Rodriguez-Aguilar, Sierra, & Vasconcelos, 2004). For the sake of clarity, in this chapter we will adopt the term “institution” to refer to a set of rules and concepts regulating agent interactions and which may be enforced by an organization (an electronic institution according to Esteva et al. (2001)). This use of the terms “institution” and “organization” is inspired by (Searle, 1995), where an institution is any collectively accepted system of rules which creates institutional facts, and by (North, 1990), where institutions are the rules of a society regulating interactions and activities of organizations.

In contrast with (Moses & Tennenholtz, 1995), where norms (also named social laws) are assumed to be respected by agents because they are designed by a single organization, in open systems it is unrealistic to expect that autonomous agents will always comply with norms. In particular, like other social concepts (e.g., expectations (Alberti, Gavanelli, Lamma, Chesani, Mello, & Torroni, 2006) or commitments (Yolum & Singh, 2004) (Fornara, Viganò, & Colombetti, 2007)), norms describe interactions in terms of public observable entities which reflect how an agent should behave. For this reason, most efforts have been devoted to the development of normative languages amenable to automatic monitoring (e.g., (García-Camino, Noriega, & Rodríguez-Aguilar, 2005)) and tools to either avoid violations (Hübner, Simão Sichman, & Boissier, 2005) or detect and sanction them at runtime (e.g., (Esteva, Rodríguez-Aguilar, Rosell, & Arcos, 2004)). Doing so increases the efficiency of transactions carried out by agents (Noriega, 1997), but raises the problem of ensuring that such systems of rules are not characterized by contradictory norms, and allow agents to fulfill the objectives of the organization that has developed the system. This is especially important when norms of institutions are complex and cannot be reduced to a simple protocol that agents should follow, making it prohibitive to foresee all possible evolutions admitted by them.

Unfortunately, most approaches do not provide tools to automatically verify what properties are guaranteed when norms are followed or violated. As a consequence, designers can specify a set of norms which can be enforced at runtime, but they cannot check if such norms are consistent, undesirable behaviors are actually forbidden, and interactions terminate with positive outcomes when norms are correctly followed. For instance, in (Viganò & Colombetti, 2007) we showed that the rules of the Dutch Auction Institution discussed in (García-Camino, Rodríguez-Aguilar, Sierra, & Vasconcelos, 2006) admit infinite bidding rounds. The lack of tools for the verification of norms defined by institutions may be problematic, since institutions have been put forward to provide safe and reliable environments for agents (Esteva, Rodriguez-Aguilar, Sierra, García, & Arcos, 2001).

Automated formal verification should be considered as an important step for the development of organizations regulated by institutions, because it can increase their reliability by ensuring that they satisfy certain properties. The development of formal frameworks to model and automatically verify organizations is therefore essential, but only few attempts have been proposed in the literature to introduce automated formal methods to verify them (e.g. (Huget, Esteva, Phelps, Sierra, & Wooldridge, 2002)). In our opinion, model checking (Clarke, Grumberg, & Peled, 1999) can play an important role to evaluate the consistency of the rules governing an agent organization, to ensure its compliance with the rules defined by a human institution, and to compare design alternatives arising from different interpretations of such rules.

The remainder of this chapter is organized as follows. In the next section we provide an overview of the main approaches proposed in the literature to model and verify systems regulated by norms. In particular, we will focus our attention on the most recent contributions in the context of multiagent