Chapter 7
The Norm Implementation Problem in Normative Multi-Agent Systems

D. Grossi, D. Gabbay, and L. van der Torre

Abstract The norm implementation problem consists in how to see to it that the agents in a system comply with the norms specified for that system by the system designer. It is part of the more general problem of how to synthesize or create norms for multi-agent systems, by, for example, highlighting the choice between regimentation and enforcement, or the punishment associated with a norm violation. In this paper we discuss how various ways to implement norms in a multi-agent system can be distinguished in a formal game-theoretic framework. In particular, we show how different types of norm implementation can all be uniformly specified and verified as types of transformations of extensive games. We introduce the notion of retarded preconditions to implement norms, and we illustrate the framework and the various ways to implement norms in the blocks world environment.
7.1 Introduction

Normative multi-agent systems (NMAS) [57] study the specification, design, and programming of systems of agents by means of systems of norms. Norms allow for the explicit specification of the standards of behavior the agents in the systems are supposed to comply with. Once such a set of norms is settled, the question arises of how to organize the agents’ interactions in the system, in such a way that those norms do not remain—so to say—dead letter, but they are actually followed by the agents. Designing a NMAS does not only mean to state a number of standards of behavior in the form of a set of norms, but also to organize the system in such a way that those standards of behavior are met by the agents participating in the system. In other words, norm creation [53] distinguishes between the creation of the obligation and norm implementation, because these two problems have different concerns. On the one hand the creation of the obligation says how the ideal can be reached, and the creation of the sanction says how agents can be motivated to comply with the norms such that the ideal will (probably) be reached. The paper moves the first steps towards a formal understanding of the norm implementation problem, defined as follows.

The norm implementation problem. How to make agents comply with a set of norms in a system?

In this paper we introduce a formal framework that can represent various solutions to the norm implementation problem, which can be used to analyze them, or to make a choice among them. For example, in some cases a norm cannot be regimented, such as the norm to return books to the library within two weeks, but in other cases there is the choice between regimentation and enforcement. It is often assumed that regimenting norms makes the system more predictable, since agents cannot violate the norms, but as a consequence it also makes the system less flexible and less efficient. Conceptually, regimentation is easier than enforcement, and since agents are bounded reasoners who can make mistakes, regimentation is often favored by policy makers. However, policy makers are bounded reasoners too, who have to make norms in uncertain circumstances, and therefore most people prefer enforcement over regimentation—at least, when the legal system is reliable. As another example, it is often assumed that very high punishments make the system less efficient than lower ones, due to the lack of incentives for agents once they have violated a norm. Such assumptions are rarely studied formally.

Although ideas about norm implementation can be found scattered over, in particular, the multi-agent systems literature (for instance, in OperA [145], Moise+ [242], AMELI [167], J-Moise+ [243], and in programming languages for multi-agent system programs [125]), they have not yet been presented in a systematic and uniform way. One of the aims of the paper is to do so, providing a formal overarching framework within which it becomes possible to place and compare existing contributions. So the first requirement on a framework for norm implementation is that it can represent existing widely discussed norm implementation methods such