

Introduction to Normative Multiagent Systems

Guido Boella¹, Leendert van der Torre² and Harko Verhagen³

¹ Dipartimento di Informatica, Università di Torino
I-10149, Torino, Corso Svizzera 185, Italy

`guido@di.unito.it`

² Department of Computer Science and Communications, University of Luxembourg
L-1359, Luxembourg, rue Richard Coudenhove 6 - Kalergi, Luxembourg

`leon.vandertorre@uni.lu`

³ Department of Computer and Systems Sciences, Stockholm University / KTH
SE-16440, Kista, Forum 100, Sweden

`verhagen@dsv.su.se`

Abstract. This article introduces the research issues related to and definition of normative multiagent systems.

Keywords. Norms, Multiagent systems, Normative multiagent systems

1 Introduction

Normative multiagent systems as a research area can be defined as the intersection of normative systems and multiagent systems. Since the use of norms is a key element of human social intelligence, norms may be essential too for artificial agents that collaborate with humans, or that are to display behavior comparable to human intelligent behavior. By integrating norms and individual intelligence normative multiagent systems provide a promising model for human and artificial agent cooperation and co-ordination, group decision making, multiagent organizations, regulated societies, electronic institutions, secure multiagent systems, and so on.

With ‘normative’ we mean ‘conforming to or based on norms’, as in *normative behavior* or *normative judgments*. According to the Merriam-Webster Online [1] Dictionary, other meanings of normative not considered here are ‘of, relating to, or determining norms or standards’, as in *normative tests*, or ‘prescribing norms’, as in *normative rules of ethics* or *normative grammar*. With ‘norm’ we mean ‘a principle of right action binding upon the members of a group and serving to guide, control, or regulate proper and acceptable behavior’. Other meanings of ‘norm’ given by the Merriam-Webster Online Dictionary but not considered here are ‘an authoritative standard or model’, ‘an average like a standard, typical pattern, widespread practice or rule in a group’, and various definitions used in mathematics.

Normative multiagent systems are an example of the use of sociological theories in multiagent systems, and more generally of the relation between agent theory and the social sciences such as sociology, philosophy, economics, and legal

science. The need for social science theories and concepts like norms in multi-agent systems is now well established. For example, Wooldridge’s weak notion of agency is based on flexible autonomous action [2], and social ability as the interaction with other agents and co-operation is one of the three meanings of flexibility; the other two are reactivity as interaction with the environment, and pro-activeness as taking the initiative. In this definition autonomy refers to non-social aspects, such as operating without the direct intervention of humans or others, and have some kind of control over their actions and internal state. For some other arguments for the need for social theory in multiagent systems, see, for example, [3,4,5]. For a more complete discussion on the need of social theory in general, and norms in particular, see the AgentLink roadmap [6].

Social concepts like norms are important for multiagent systems, because multiagent system research and sociology share the interest in the relation between micro-level agent behaviour and macro-level system effects. In sociology this is the (in)famous micro-macro link [7] that focuses on the relation between individual agent behaviour and characteristics at the level of the social system. In multiagent system research, this boils down to the question “How to ensure efficiency at the level of the multiagent system whilst respecting individual autonomy?”. According to Verhagen [8] three possible solutions to this problem comprise of the use of central control which gravely jeopardizes the agent’s autonomy, internalized control like the use of social laws [9], and structural co-ordination [10] including learning norms.

Before we discuss normative multiagent systems, we consider some discussions on norms in the social sciences.

2 Norms and normative systems

In the 1960’s, the sociologist Gibbs [11] wrote an influential article on the problems concerning the definition and classification of norms, and observes that the various types of norms involve “a collective evaluation of behavior in terms of what it *ought* to be; a collective expectation as to what behavior *will be*; and/or particular *reactions* to behavior, including attempts to apply sanctions or otherwise induce a particular kind of conduct.” [11, p. 589, original emphasis]

More recently, Therborn [12] presented an overview of the role of norms for social theory and analysis. Normative action is based upon wanting to do the right thing rather than the thing that leads to ends or goals, which he calls teleological action, or the thing that leads to, expresses, or is caused by an emotion, called emotional action.

Therborn distinguishes among three kinds of norms. *Constitutive norms* define a system of action and an agent’s membership in it, *regulative norms* describe the expected contributions to the social system, and *distributive norms* defining how rewards, costs, and risks are allocated within a social system. Furthermore, he distinguishes between non-institutionalized normative order, made up by personal and moral norms in day-to-day social traffic, and institutions, an example of a social system defined as a closed system of norms. Institutional normative

action is equaled with role plays, i.e., roles find their expressions in expectations, obligations, and rights vis-a-vis the role holder's behaviour.

Therborn also addresses the dynamics and changing of norms. The dynamics of norms at the level of the individual agent is how norms are learned or propagated in a population. Socialization is based on identification, perceiving the compliance with the norms by other agents, or the entering of an institution. Norms are (re)enforced by the presence of incentives or sanctions. Changes in either of these three socialization mechanisms lead to changes in the set of norms of the individual agent. These changes may be inhibited either by changes in the social system or changed circumstances, or by changes in the interpretation of the norms by the agents within the system.

Within philosophy normative systems have traditionally been studied by moral and legal philosophers. Alchourròn and Bulygin [13] argue that a normative system should not be defined as a set of norms, as is commonly done, but in terms of consequences:

“When a deductive correlation is such that the first sentence of the ordered pair is a case and the second is a solution, it will be called normative. If among the deductive correlations of the set α there is at least one normative correlation, we shall say that the set α has normative consequences. A system of sentences which has some normative consequences will be called a normative system.” [13, p.55].

In computer science, Meyer and Wieringa define normative systems as “systems in the behavior of which norms play a role and which need normative concepts in order to be described or specified” [14, preface]. They also explain why normative systems are intimately related with deontic logic.

“Until recently in specifications of systems in computational environments the distinction between normative behavior (as it *should be*) and actual behavior (as it *is*) has been disregarded: mostly it is not possible to specify that some system behavior is non-normative (illegal) but nevertheless possible. Often illegal behavior is just ruled out by specification, although it is very important to be able to specify what should happen if such illegal but possible behaviors occurs! Deontic logic provides a means to do just this by using special modal operators that indicate the status of behavior: that is whether it is legal (normative) or not” [14, preface].

3 Normative multiagent systems

The agents in the environment of a normative system interact with the normative system in various ways. First, from the perspective of the agents, agents can create new norms, update or maintain norms, and enforce norms, using roles defined in the normative system such as legislators or policemen. Secondly, from the perspective of social order, we can also look at the interaction between the normative system and its environment from the viewpoint of the normative

system. In this viewpoint, the normative system uses the agents playing a role in it – the legislators, policemen and the like – to maintain an equilibrium in the normative multiagent system. In this perspective, we can distinguish at least two levels of equilibrium. First, norms are used to maintain social order in a normative multiagent system. Second, normative system contain a mechanism for updating themselves, to adapt to changing circumstances in its environment.

Jones and Carmo [15] define a normative system as “Sets of agents whose interactions are norm-governed; the norms prescribe how the agents ideally should and should not behave. [...] Importantly, the norms allow for the possibility that actual behavior may at times deviate from the ideal, i.e., that violations of obligations, or of agents’ rights, may occur.” In our opinion, this is too general, as a normative system does not contain the agents themselves. It also is not a satisfactory definition of normative multiagent system, because it precludes the agents’ control over the set of norms. We therefore use the following definition in this paper.

A normative multiagent system is a multiagent system together with normative systems in which agents on the one hand can decide whether to follow the explicitly represented norms, and on the other the normative systems specify how and in which extent the agents can modify the norms.

Note that this definition makes no presumptions about the internal workings of an agent nor of the way norms find their expression in agent’s behaviour.

Since norms are explicitly represented, according to our definition of a normative multiagent system, the question should be raised how norms are represented. Norms can be interpreted as a special kind of constraint, and represented depending on the domain in which they occur. However, the representation of norms by domain dependent constraints runs into the question what happens when norms are violated. Not all agents behave according to the norm, and the system has to deal with it. In other words, norms are not hard constraints, but soft constraints. For example, the system may sanction violations or reward good behavior. Thus, the normative system has to monitor the behavior of agents and enforce the sanctions. Also, when norms are represented as domain dependent constraints, the question will be raised how to represent permissive norms, and how they relate to obligations. Whereas obligations and prohibitions can be represented as constraints, this does not seem to hold for permissions. For example, how to represent the permission to access a resource under an access control system? Finally, when norms are represented as domain dependent constraints, the question can be raised how norms evolve.

We therefore believe that norms should be represented as a domain independent theory, for example in deontic logic [16,17,18,19,20,21]. Deontic logic studies logical relations among obligations and permissions, and more in particular violations and contrary-to-duty obligations, permissions and their relation to obligations, and the dynamics of obligations over time. Therefore, insights from deontic logic can be used to represent and reason with norms. Deontic logic also offers representations of norms as rules or conditionals. However, there

are several aspects of norms which are not covered by constraints nor by deontic logic, such as the relation between the cognitive abilities of agents and the global properties of norms.

Conte, Falconi and Sartor [22] say that normative multiagent systems research focuses on two different sets of problems. On the one hand, they claim that legal theory and deontic logic supply a theory for of norm-governed interaction of autonomous agents while at the same time lacking a model that integrates the different social and normative concepts of this theory. On the other hand, they claim that three other problems are of interest in multiagents systems research on norms: how agents can acquire norms, how agents can violate norms, and how an agent can be autonomous. For artificial agents, norms can be designed as in legal human systems, forced upon, for example when joining an institution, or they can emerge from the agents making them norm autonomous [8]. Agent decision making in normative systems and the relation between desires and obligations has been studied in agent architectures [23], which thus explain how norms and obligations influence agent behavior.

An important question is where norms come from. Norms are not necessarily created by a single legislator, they can also emerge spontaneously, or be negotiated among the agents. In electronic commerce research, for example, cognitive foundations of social norms and contracts are studied [24]. Protocols and social mechanisms are now being developed to support such creations of norms in multiagent systems. When norms are created, the question how they are enforced can be raised. For example, when a contract is violated, the violator may have to pay a penalty. But then there has to be a monitoring and sanctioning system, for example police agents in an electronic institution. Such protocols or roles in a multiagent system are part of the construction of social reality, and Searle [25] has argued that such social realities are constructed by constitutive norms. This again raises the question how to represent such constitutive or counts-as norms, and how they are related to regulative norms like obligations and permissions [24].

Not only the relation between norms and agents must be studied, but also the relation between norms and other social and legal concepts. How do norms structure organizations? How do norms coordinate groups and societies? How about the contract frames in which contracts live? How about the legal contexts in which contract frames live? How about the relation between legal courts? Though in some normative multiagent systems there is only a single normative system, there can also be several of them, raising the question how normative systems interact. For example, in a virtual community of resource providers each provider may have its own normative system, which raises the question how one system can authorize access in another system, or how global policies can be defined to regulate these local policies [26].

Summarizing, normative multiagent systems study general and domain independent properties of norms. It builds on results obtained in deontic logic, the logic of obligations and permissions, for the representation of norms as rules, the application of such rules, contrary-to-duty reasoning and the relation to per-

missions. However, it goes beyond logical relations among obligations and permissions by explaining the relation among social norms and obligations, relating regulative norms to constitutive norms, explaining the evolution of normative systems, and much more.

Some of these issues can be discussed in more detail. These include action (e.g., the BDI model of agency) with models of normative action, to be combined, reasoning and dynamics, and theories of normative action into implementable formal models.

General themes that are to be addressed in research on normative agent systems include

1. intra-agent aspects of norms,
2. interagent aspects of norms,
3. normative systems and their borders, and
4. combining normative systems.

In [27] a collection of articles on these issues is presented.

References

1. Merriam-Webster OnLine: The Language Center (2005) www.m-w.com/.
2. Wooldridge, M.: An Introduction to MultiAgent Systems. Wiley (2002)
3. Bond, A.H., Gasser, L.: An Analysis of Problems and Research in DAI. In Bond, A.H., Gasser, L., eds.: Readings in Distributed Artificial Intelligence, Morgan Kaufmann (1988) 3–35
4. Conte, R., Gilbert, N.: Computer Simulation for Social Theory. In: Computer Simulation for Social Theory. UCL Press. (1995) 1 – 18
5. Verhagen, H., Smit, R.: Modelling Social Agents in a Multiagent World. In Van de Velde, W., Perram, J.W., eds.: , Position Papers MAAMAW 1996, Technical Report 96-1, Vrije Universiteit Brussel - Artificial Intelligence Laboratory (1996)
6. : Agent Technology Roadmap: A Roadmap for Agent-Based Computing. (2005)
7. Alexander, J., Giesen, B., Münch, R., Smelser, N., eds.: The Micro-Macro Link. University of California Press (1987)
8. Verhagen, H.: Norm Autonomous Agents. PhD thesis, Department of System and Computer Sciences, The Royal Institute of Technology and Stockholm University, Sweden (2000)
9. Shoham, Y., Tennenholtz, M.: On the Synthesis of Useful Social Laws for Artificial Agent Societies (Preliminary Report). In: Proceedings of the National Conference on Artificial Intelligence, San Jose, CA (1992) 276–281
10. Ossowski, S.: Co-ordination in Artificial Agent Societies. Springer (1999)
11. Gibbs, J.P.: Norms: The Problem of Definition and Classification. The American Journal of Sociology **70** (1965) 586 – 594
12. Therborn, G.: Back to Norms! On the Scope and Dynamics of Norms and Normative Action. Current Sociology **50** (2002) 863 – 880
13. Alchourrón, C., Bulygin, E.: Normative Systems. Springer (1971)
14. Meyer, J.J., Wieringa, R., eds.: Deontic Logic in Computer Science: Normative System Specification. Wiley (1993)
15. Jones, A., Carmo, J.: Deontic Logic and Contrary-to-Duties. In Gabbay, D., ed.: Handbook of Philosophical Logic. Kluwer (2001) 203279

16. von Wright, G.: *Mind*. **60** (1951) 1–15
17. van der Torre, L., Tan, Y.: Contrary-to-duty reasoning with preference-based dyadic obligations. *Annals of Mathematics and Artificial Intelligence* **27** (1999) 49–78
18. van der Torre, L.: Contextual deontic logic: Normative agents, violations and independence. *Annals of Mathematics and Artificial Intelligence* **37(1-2)** (2003) 33–63
19. Makinson, D., van der Torre, L.: Input-output logics. *Journal of Philosophical Logic* **29** (2000) 383–408
20. Makinson, D., van der Torre, L.: Constraints for input-output logics. *Journal of Philosophical Logic* **30(2)** (2001) 155–185
21. Makinson, D., van der Torre, L.: Permissions from an input/output perspective. *Journal of Philosophical Logic* **32(4)** (2003) 391–416
22. Conte, R., Falcone, R., Sartor, G.: Introduction: Agents and Norms: How to Fill the Gap? *Artificial Intelligence and Law* (1999) 1 – 15
23. Broersen, J., Dastani, M., Hulstijn, J., van der Torre, L.: Goal generation in the BOID architecture. *Cognitive Science Quarterly* **2(3-4)** (2002) 428–447
24. Boella, G., van der Torre, L.: A game theoretic approach to contracts in multiagent systems. *IEEE Trans. SMC, Part C* (2006)
25. Searle, J.R.: *The Construction of Social Reality*. The Free Press (1995)
26. Boella, G., van der Torre, L.: Security policies for sharing knowledge in virtual communities. *IEEE Trans. SMC, Part A* (2006)
27. Boella, G., Van der Torre, L., Verhagen, H., eds.: Special Issue on Normative Multiagent Systems - *Journal of Computational & Mathematical Organization Theory*. Volume 12 (2 - 3). (2006)