Introduction to Normative Multiagent Systems

Guido Boella
Dipartimento di Informatica
Università di Torino
Italy
guido@di.unito.it

Leendert van der Torre
Department of Computer Science
University of Luxembourg
Luxembourg
leendert@vandertorre.com

Harko Verhagen
Dept. of Computer and Systems Sciences
Stockholm University / KTH,
Forum 100, SE-16440 Kista, Sweden
verhagen@dsv.su.se

May 29, 2006

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 de-

cision 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 (2005) 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 multiagent systems is now well established. For example, Wooldridge's weak notion of agency is based on flexible autonomous action (Wooldridge, 2002), 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, (Bond and Gasser, 1988; Conte and Gilbert, 1995; Verhagen and Smit, 1996). For a more complete discussion on the need of social theory in general, and norms in particular, see the AgentLink roadmap (roa, 2005).

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 (Alexander et al., 1987) 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 (2000) 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 (Shoham and Tennenholtz, 1992), and structural coordination (Ossowski, 1999) including learning norms.

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

1 Norms and normative systems

In the 1960's, the sociologist Gibbs (1965) 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." (Gibbs, 1965, p. 589, original emphasis)

More recently, Therborn (2002) 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 sys-

tem 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 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 (1971) 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." (Alchourròn and Bulygin, 1971, 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" (Meyer and Wieringa, 1993, 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" (Meyer and Wieringa, 1993, preface).

2 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 (2001) 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 (von Wright, 1951; van der Torre and Tan, 1999; van der Torre, 2003; Makinson and van der Torre, 2000; Makinson and van der Torre, 2001; Makinson and van der Torre, 2003). 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 (1999) 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 (Verhagen, 2000). Agent decision making in normative systems and the relation between desires and obligations has been studied in agent architectures (Broersen et al., 2002), 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 (Boella and van der Torre, 2006a). 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 (1995) 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 (Boella and van der Torre, 2006a).

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 (Boella and van der Torre, 2006b).

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 permissions. 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.

The papers in this double special issue on normative multiagent systems address some of these issues, but they also address new research issues that are of central importance for the whole field of normative multiagent systems. These include how to combine theories of teleological action (e.g., the BDI model of agency) with models of normative action, how to model the dynamics of norms when institutions' norm sets are to be combined, the development and testing of logics of normative reasoning and dynamics, and the formalization descriptive social theories of normative action into implementable formal models.

3 NorMAS 2005

NorMAS05 was an international symposium on normative multiagent systems, organized in April 2005 by the authors of this article as part of the 2005 AISB convention (AISB standing for the Society for the Study of Artificial Intelligence and the Simulation of Behaviour). The symposium attracted papers from a variety of areas, such as the social sciences (and computational sociology in particular), computer science, and formal logics. A number of these papers representing these areas were selected for this double special issue on normative multiagent systems. Four general themes are addressed in these papers, namely intra-agent aspects of norms, interagent aspects of norms, normative systems and their borders, and combining normative systems.

3.1 Intra-agent aspects of norms

The paper "My Agents Love to Conform: Norms and Emotion in the Micro-Macro Link" by von Scheve et al. investigates the function of emotion in relation to norms in natural and artificial societies. It shows that unintentional behavior can be normative and socially

functional at the same time, thereby highlighting the role of emotion. By defining norms as mental objects, the role of emotion in maintaining and enforcing norms is studied, relates these findings social structural dynamics in natural and societies,and outlines the possibilities of an application to a multi-agent architecture.

Sadri, Stati, and Toni's "Normative KGP Agents" extends the logical model of agency known as the KGP model to support agents with normative concepts, based on the roles an agent plays and the obligations and prohibitions that result from playing these roles. The proposed framework illustrates how the resulting normative concepts, including the roles, can evolve dynamically during the lifetime of the agent. It also illustrates how these concepts can be combined with the existing capabilities of KGP agents in order to plan for their goals, react to changes in the environment, and interact with other agents. Finally, the paper gives an executable specification of normative concepts that can be used directly for prototyping applications.

3.2 Interagent aspects of norms

Kibble's paper "Speech acts, commitment and multiagent communication" aims to reconsider the suitability of speech act theory as a basis for agent communication languages. It models dialogue states as deontic scoreboards which keep track of commitments and entitlements that speakers acknowledge and hearers attribute to other interlocutors and outlines an update semantics and protocol for selected locutions.

Sauro's paper "Qualitative Criteria of Admissibility for Enforced Agreements" focuses on the desirablility of artificial agents to help each other when they cannot achieve their goals, or when they profit from social exchanges. It studies the coalition formation processes supported by enforced agreements and defines two qualitative criteria that establish when a coalition is admissible to be formed. These two properties can be used when the space of possible coalitions is unknown.

3.3 Normative systems and their borders

Davidsson and Johansson classify artificial societies and identify four different types of stakeholders in their paper "On the Potential of Norm-Governed Behavior in Different Categories of Artificial Societies". The potential of norm-governed behavior in different types of artificial societies is investigated based on the preferences of the stakeholders and how they influence the state of the society. The paper concludes that the more open a society is the more it has to rely on agent owners and designers to achieve norm-governed behavior, whereas in more closed societies the environment designers and owners may control the degree of norm-governed behavior.

Hahn, Fley, and Florian argue in "A Framework for the Design of Self-Regulation of Open Agent-based Electronic Marketplaces" that allowing self-interested agents to activate social institutions during run-time can improve the robustness of open multiagent systems. Based on sociological theory, institutions are seen as rules which have to be activated and adopted by the agent population. A framework for self-regulation of multiagent system for the domain of electronic marketplaces is developed, consisting of three different institutional forms that are defined by the mechanisms and instances that generate, change, or safeguard them. The paper shows that allowing autonomous agents both the reasoning about their compliance with a rule and the selection of the form of an institution helps to balance the trade-off between the autonomy of self-interested agents and the maintenance of social order in an open multiagent system and to ensure almost the same qualities as in closed environments.

In "Mapping Deontic Operators to Abductive Expectations", Alberti et al. propose a mapping of deontic operators (obligations, prohibition, permission) to language entities (expectations) available within the an agent framework developed for agent interaction in open agent societies. The mapping is supported by showing a similarity between the abductive semantics for expectations and the Kripke semantics that can be given to deontic

operators.

In "A Normative Framework for Agent-Based Systems", López y López, Luck, and d'Inverno present a formal normative framework for agent-based systems that adresses two omissions of previous research on the use of norms in computational models of open societies to help to cope with the heterogeneity, the autonomy and the diversity of interests among their members. These are the lack of a canonical model of norms that facilitates their implementation and enables the description of the processes of reasoning about norms, and secondly the perspective of individual agents and what they might need to effectively reason about the society in which they participate.

3.4 Combining normative systems

Grossi et al. introduce the notion of contextual ontologies in their paper "Ontological Aspects of the Implementation of Norms in Agent-Based Electronic Institutions" and also provide a formal machinery to characterise this notion. This notion solves the problem of different institutions implementing the same set of norms in different ways presupposing divergent ontologies of the concepts in which that set of norms is formulated.

References

2005. *Agent Technology Roadmap: A Roadmap for Agent-Based Computing.*

Alchourròn, C.E. and E. Bulygin. 1971. *Normative Systems.* Springer.

Alexander, J.C., B. Giesen, R. Münch, and N.J. Smelser, editors. 1987. *The Micro-Macro Link.* University of California Press.

Boella, G. and L. van der Torre. 2006a. A game theoretic approach to contracts in multiagent systems. *IEEE Trans. SMC, Part C.*

Boella, G. and L. van der Torre. 2006b. Security policies for sharing knowledge in virtual communities. *IEEE Trans. SMC, Part A*.

Bond, A. H. and L. Gasser. 1988. An Analysis of Problems and Research in DAI. In A. H. Bond and L. Gasser, editors, *Readings in Distributed Artificial Intelligence*, pages 3–35. Morgan Kaufmann.

Broersen, J., M. Dastani, J. Hulstijn, and L. van der Torre. 2002. Goal generation in the BOID architecture. *Cognitive Science Quarterly*, 2(3-4):428–447.

Conte, R., R. Falcone, and G. Sartor. 1999. Introduction: Agents and Norms: How to Fill the Gap? *Artificial Intelligence and Law*, pages 1 – 15.

Conte, R. and N. Gilbert, 1995. "Computer Simulation for Social Theory". In R. Conte and N. Gilbert, editors, *Artificial Societies: The Computer Simulation of Social Life*, chapter Computer Simulation for Social Theory, pages 1 – 18. UCL Press.

Gibbs, J. P. 1965. Norms: The Problem of Definition and Classification. *The American Journal of Sociology*, 70(5):586 – 594.

Jones, A. and J. Carmo. 2001. Deontic Logic and Contrary-to-Duties. In D. Gabbay, editor, *Handbook of Philosophical Logic*. Kluwer, page 203279.

Makinson, D. and L. van der Torre. 2000. Input-output logics. *Journal of Philosophical Logic*, 29:383–408.

Makinson, D. and L. van der Torre. 2001. Constraints for input-output logics. *Journal of Philosophical Logic*, 30(2):155–185.

Makinson, D. and L. van der Torre. 2003. Permissions from an input/output perspective. *Journal of Philosophical Logic*, 32(4):391–416.

Merriam-Webster OnLine. 2005. The Language Center. `www.m-w.com/`.

Meyer, J-J. and R. Wieringa, editors. 1993. *Deontic Logic in Computer Science: Normative System Specification*. Wiley.

Ossowski, S. 1999. *Co-ordination in Artificial Agent Societies*. Springer.

Searle, J. R. 1995. *The Construction of Social Reality*. The Free Press.

Shoham, Y. and M. Tennenholtz. 1992. On the Synthesis of Useful Social Laws for Artificial Agent Societies (Preliminary Report). In *Proceedings of the National Conference on Artificial Intelligence*, pages 276–281, San Jose, CA.

Therborn, G. 2002. Back to Norms! On the Scope and Dynamics of Norms and Normative Action. *Current Sociology*, 50(6):863 – 880.

van der Torre, L. 2003. Contextual deontic logic: Normative agents, violations and independence. *Annals of Mathematics and Artificial Intelligence*, 37(1-2):33–63.

van der Torre, L. and Y. Tan. 1999. Contrary-to-duty reasoning with preference-based dyadic obligations. *Annals of Mathematics and Artificial Intelligence*, 27:49–78.

Verhagen, H. 2000. *Norm Autonomous Agents*. Ph.D. thesis, Department of System and Computer Sciences, The Royal Institute of Technology and Stockholm University, Sweden.

Verhagen, H. and R. Smit. 1996. Modelling Social Agents in a Multiagent World. In W. Van de Velde and J. W. Perram, editors, *, Position Papers MAAMAW 1996, Technical Report 96-1*. Vrije Universiteit Brussel - Artificial Intelligence Laboratory.

von Wright, G.H. 1951. Deontic logic. 60:1–15.

Wooldridge, M. 2002. *An Introduction to MultiAgent Systems*. Wiley.

Bibliography

Guido Boella

Guido Boella received the PhD degree at the University of Torino in 2000. He is currently professor at the Department of Computer Science of the University of Torino. His research interests include multi-agent systems, in particular, normative systems, institutions and

roles using qualitative decision theory. He is the co-chair of the first workshops on normative multi-agent systems (NorMas05), on coordination and organization (CoOrg05), and the AAAI Fall Symposium on roles (Roles05).

Leendert van der Torre

Leendert van der Torre received the Ph.D. degree in computer science from Erasmus University Rotterdam, The Netherlands, in 1997. He is currently a Full Professor at the University of Luxembourg. He has developed the so-called input/output logics and the BOID agent architecture. His current research interests include deontic logic, qualitative game theory, and security and coordination in normative multiagent systems.

Harko Verhagen

Harko Verhagen received his Ph.D. degree in computer and systems sciences from Stockholm University (Sweden) in 2000 and is currently an associate professor at the department. His research has focussed on simulation of organizational behaviour, simulation as a scientific method, the use of sociological theories in multiagent systems research and more in particular theories on norms and autonomy.