Decentralized Control
Obligations and permissions in virtual communities of agents

Guido Boella[1] and Leendert van der Torre[2]

[1] Dipartimento di Informatica, Università di Torino, Italy
guido@di.unito.it
[2] CWI, Amsterdam, The Netherlands
torre@cwi.nl

Abstract. In this paper we introduce a model of local and global control policies regulating decentralized virtual communities of heterogeneous agents. We illustrate how the model can be formalized if agents attribute mental attitudes to normative systems.

1 Introduction

In highly decentralized structures such as peer-to-peer systems and grid architectures, there is neither a central design nor a central administrator, agents play both the role of resource consumers and providers, and they have the power to control their resources by expressing their own *local access policies*. But to form a *virtual community* local access policies should be organized according to global policies that define resource sharing among the members of the community. According to Pearlman *et al.* [10], "a key problem associated with the formation and operation of distributed virtual communities is that of how to specify and enforce community policies." Since there is no plausible way to enforce the respect of global policies by constraining the architecture, it is necessary to have a *normative system* able to specify global policies about local policies.

In [2] we argue that the interaction between an agent and its normative system can be modelled as a game between them when the agent attributes mental attitudes to the normative system, and in [3] we extended this argument to norm creation. In this paper we address the following two questions:

1. How to model decentralized control in virtual communities using a normative system with local and global control policies?
2. How can the attribution of mental attitudes to normative systems be used to reason about obligations and permissions regulating virtual communities of heterogeneous agents?

This paper is organized as follows. In Section 2 we discuss the first question of modelling decentralized control, in Section 3 we discuss the second question on the attribution of mental attitudes to normative systems, and in Section 4 we discuss how our model of decentralized control could be formalized.

2 How to model decentralized control?

A normative system which specifies what is obliged or permitted is not sufficient to regulate a virtual community, because the "exercise of rights is effective only if the resource provider has granted those rights to the community" [10]. Moreover, the global authority must be able to *motivate* local authorities such as these resource providers to issue policies which respect the global ones, besides motivating resource consumers to respect the norms. However, Firozabadi and Sergot [9] argue that in a virtual community of heterogeneous agents, a local authority can deliberately fail to comply with global policies: "Upon an access request from a_1, a_2 has to decide whether to grant the access to a_1 or not. There are several possible cases:

1. a_1 is permitted to access the resource, but there is no obligation on a_2 to grant that access. a_2 will not violate any policy regardless of whether it grants or denies the access.
2. a_1 is not only permitted to access the resource, but is also *entitled* to it. This means that a_2 has an obligation to grant the access whenever a_1 requests it. A typical scenario is when a_1 is the owner of d and a_2 is the storage service provider for d.
3. a_1 has no permission to access d, and so a_2 is forbidden to grant the access. Note that a_2 may have the practical possibility to give access to a_1 even if it is not permitted to do so."

Consequently, a model of community policies and decentralized control should be able to deal with the following two issues. First, it should model local authorities that can deliberately fail to comply with global policies, and it should model how the global authority motivates and sanctions local authorities. In [2] we argue that a model of norm-evading agents can be based on interpreting normative systems as agents, for example by attributing mental attitudes to them. This attribution of mental attitudes is also the basis of the model in this paper.

Second, it should model various distinctions. E.g., Firozabadi and Sergot observe that in the literature on computer security, the terms *right*, *permission* and *privilege* often have the same meaning, whereas the term *entitlement* emphasizes a concept which is stronger than mere permission. In this paper, notions like entitlement are expressed as the obligations and permissions posed by a global authority on access permissions of the resource consumers. Entitlement is an obligation of the global authority to make a behavior permitted. Moreover, since something is obliged if its absence counts as a violation, it is defined as the obligation that the local authority does not count this behavior as a violation (2. above). The local authority, however, can still violate this global policy and forbid access to users if it prefers to face the sanction. Similarly, the obligation to forbid access can be modelled as a global obligation that local authorities count access as a violation (3. above). The permission to forbid access never leads to a violation of the local authority, regardless of whether it counts access as a violation. Analogously, the permission to permit access does not lead to a violation if the local authority does not count access as a violation (1. above).

3 Our approach: normative systems as agents!

Normative systems that control and regulate behavior like legal, moral or security systems are autonomous, they react to changes in their environment, and they are pro-active. For example, the process of deciding whether behavior counts as a violation is an autonomous activity. Since these properties have been identified as the properties of autonomous or intelligent agents by [11], normative systems may be called *normative agents*. This goes beyond the observation that a normative system may contain agents, like a legal system contains legislators, judges and policemen, because *a normative system itself is called an agent*.

The first advantage of the normative systems as agents perspective is that the interaction between an agent and the normative system which creates and controls its obligations can be modelled as a game between two agents. Consequently, methods and tools used in game theory such as equilibrium analysis can be applied to normative reasoning. For example, the game theories in [1, 2] are based on *recursive modelling* of the normative system by the bearer of the obligation. The agent bases its decision on the consequences of the normative system's anticipated reaction, in particular, whether the system considers the agent a violator and thus sanctions it. Analogously, the normative system can base its decision regarding which norm to create on the consequences of the agent's anticipated reaction [3].

The second advantage of the normative systems as agents perspective is that, since mental attitudes can be attributed to agents, we can attribute mental attitudes to normative systems. A consequence of the second advantage is that obligations can be defined in the standard BDI framework. In particular, Boella and Lesmo [1] suggest that we can attribute mental attitudes to normative systems, such that obligations of an agent can be interpreted as the wishes, desires or goals of the normative system. The motivation of their interpretation is the study of reasons why agents fulfil or violate sanction-based obligations.

In this paper, normative systems do not contain one authority only but they are composed of a set of authorities which are in control of their own resources. A distinguished authority (usually called *community authorization service*) plays the role of a global authority which, even if it is not in control of the local resources, issues the norms representing local policies and negotiates the conditions for the participation of agents to the virtual community. We model global policies which forbid or permit local policies to constrain the behavior of users by norms concerning other norms.

The global authority creates obligations about which behaviors the local authorities should count as violations and thus sanction. This is its way to motivate local authorities. There are two ways to formalize the relation between global and local authorities when obligations are interpreted as the goals of the authority: either the global authority has the goal that the local authority has the goal to grant access, or the global authority has the goal that the local authority grants access, but it not necessarily has the goal that the local authority has this goal. In this paper we choose the second solution, since the local authorities cannot be assumed to have the same goal as the global one.

4 Toward formalization

The agents' abilities, their beliefs and their motivations (goals and desires) must be distinguished. For example, these mental attitudes can be modelled as conditional rules in a qualitative decision theory inspired by the BOID architecture [7]. Belief rules can be used to infer the beliefs of agents using a priority relation to resolve conflicts. Goal and desire rules can be used to value a decision according to which motivations remain unsatisfied. Moreover, it must be formalized when behavior counts as a violation. There are two distinctions with the formalization sketched in [2, 3].

- In [2, 3] an obligation for x is defined as the belief that absence of a counts as a violation of some norm n. In this paper, we do not explicitly formalize the norm n. Instead, we write $V(\neg x)$ for 'the absence of x counts as a violation', see [6] for further details on this approach.
- In contrast to [2, 3] the counts-as-a-violation operator is indexed by an agent, we write $V_A(\neg x)$ for 'the absence of x counts as a violation by agent A'. Moreover, we also consider nested counts-as-a-violations. For example, we write $V_B(\neg V_A(\neg x))$ for 'the absence of counting x as a violation by agent A counts as a violation by agent B'.

Let agent A be a resource consumer agent B a resource provider, and consider the perspective of agent A. If agent A is obliged to x in context q, then agent B may decide that the absence of x counts as a violation and that agent A must be sanctioned with s, written as $O_{A,B}(x, s \mid q)$, if and only if:

1. Agent A believes that agent B desires that agent A does x if q.
2. Agent A believes that agent B desires $\neg V_A(\neg x)$, that there is no violation by agent A, but if agent B believes $\neg x \wedge q$ then it has the goal $V_A(\neg x)$, $\neg x$ counts as a violation by agent A.
3. Agent A believes that agent B desires $\neg s$, not to sanction, but if agent B decides that x is a violation by agent A, $V_A(\neg x)$, then it has as a goal that it sanctions agent A by doing s. Agent B only sanctions in case of violation. Moreover, agent A believes that agent B has a way to apply the sanction.
4. Agent A desires $\neg s$: it does not like the sanction.

Symmetrically, permission can be modelled as an exceptional situation which does not count as a violation. A permission to do $\neg x$ in context q, $P_{A,B}(\neg x \mid q)$ is an exception to an obligation to do x if in context q, if and only if agent B has the goal that $\neg x$ does not count as a violation if q.

Finally, entitlement and other nested concepts are defined as follows, where agent C is a global authority:

- Agent B believes that it is *obliged* by agent C to oblige agent A to do x with sanction s in context q iff $O_{B,C}(V_A(\neg x), s|q \wedge \neg x)$;
- Agent B believes that it is *obliged* by agent C to permit agent A not to do x with sanction s in context q iff $O_{B,C}(\neg V_A(\neg x), s'|q \wedge \neg x)$;
- Agent B believes that it is *permitted* by agent C to permit agent A not to do x in context q iff $P_{B,C}(\neg V_A(\neg x)|q \wedge \neg x)$.

5 Concluding remarks

A consequence of the attribution of mental attitudes to normative systems [2, 3] is that the multi agent system and the normative system are unified into a single system. In this paper we show that this unification is useful to model decentralized control, because also in virtual communities the distinction between multi agent system and normative system is disappearing.

In related research we further detail our model of decentralized control in virtual communities by introducing the defender role in obligations [4] and studying decisions of agents reasoning about local and global policies [5]. Further issues of research are the distinction of the roles of legislative authorities (creating norms), judicial (deciding if behavior counts as a violation) and executive ones (applying sanctions), where the former can be based on the creation of norms developed in [3]. Another issue for further research is the distinction between enacting a permission and granting an authorization [8].

References

1. G. Boella and L. Lesmo. A game theoretic approach to norms. *Cognitive Science Quarterly*, 2(3-4):492–512, 2002.
2. G. Boella and L. van der Torre. Attributing mental attitudes to normative systems. In *Proceedings of the Second International Joint Conference on Autonomous Agents and Multi Agent Systems (AAMAS'03)*, 2003.
3. G. Boella and L. van der Torre. Rational norm creation: Attributing mental states to normative systems, part 2. In *Proceedings of the Eighth International Conference on Artificial Intelligence and Law (ICAIL'03)*. ACM, 2003.
4. G. Boella and L. van der Torre. Norm governed multiagent systems: The delegation of control to autonomous agents. In *Proceedings of the 2003 IEEE/WIC International Conference on Intelligent Agent Technology (IAT'03)*. IEEE, 2003.
5. G. Boella and L. van der Torre. Local policies for the control of virtual communities. In *Proceedings of the 2003 IEEE/WIC International Conference on Web Intelligence (WI'03)*. IEEE, 2003.
6. G. Boella and L. van der Torre. Obligations as social constructs. In *Proceedings of the Italian Conference on Artificial Intelligence (AI*IA'03)*, LNAI. Springer, 2003.
7. J. Broersen, M. Dastani, J. Hulstijn, and L. van der Torre. Goal generation in the BOID architecture. *Cognitive Science Quarterly*, 2(3-4):428–447, 2002.
8. E. Bulygin. Permissive norms and normative systems. In A. Martino and F. S. Natali, editors, *Automated Analysis of Legal Texts*, pages 211–218. Publishing Company, Amsterdam, 1986.
9. B. S. Firozabadi and M. Sergot. Contractual access control. In *Proceedings of the Tenth International Workshop of Security Protocols*, Cambridge (UK), 2002.
10. L. Pearlman, V. Welch, I. Foster, C. Kesselman, and S. Tuecke. A community authorization service for group collaboration. In *Proceedings of the Third International Workshop on Policies for Distributed Systems and Networks*, pages 50–59. IEEE, 2002.
11. M. J. Wooldridge and N. R. Jennings. Intelligent agents: Theory and practice. *Knowledge Engineering Review*, 10(2):115–152, 1995.