Reasoning About Norm Compliance

(Extended Abstract)

N. Criado
Universidad Politécnica de Valencia
Camino de Vera, s/n. 46022.
Valencia, Spain
ncriado@dsic.upv.es

E. Argente
Universidad Politécnica de Valencia
Camino de Vera, s/n. 46022.
Valencia, Spain
eargente@dsic.upv.es

V. Botti
Universidad Politécnica de Valencia
Camino de Vera, s/n. 46022.
Valencia, Spain
vbotti@dsic.upv.es

P. Noriega
IIIA-CSIC
Campus de la UAB, Bellaterra,
Catalonia (Spain)
pablo@iiia.csic.es

ABSTRACT

This paper proposes a reasoning process to allow agents to decide when and how norms should be violated or obeyed. The coherence-based reasoning mechanism proposed in this paper, allows *norm aware* agents to confront the norm compliance dilemma and build alternatives for such normative decisions.

Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]: Intelligent agents

General Terms

Theory

Keywords

Norm compliance, Coherence, BDI agents

1. INTRODUCTION

The conventional understanding of regulated open MAS presumes the existence of autonomous rational agents that are subject to some explicit conventions that regulate their behaviour. Of special interest are those systems where conventions may be understood as norms and agents may decide whether to comply with those that are in force at any given time. In this paper we look into that problem, not from the normative system designer's perspective but from that of the individual agent who faces the dilemma. We propose an architecture for agents whose deliberations are aware of those norms that currently apply to them.

The main topic addressed by this paper is the problem of making decisions about violating or obeying norms. Specifically, a reasoning process for making decisions about norm compliance is proposed. This mechanism has been applied in a Normative BDI Architecture (or n-BDI for short) [2]. The n-BDI proposal is an extension of a Multi-Context Graded BDI architecture [1] with an explicit representation of norms.

Cite as: Reasoning About Norm Compliance (Extended Abstract), N. Criado, E. Argente, V. Botti and P. Noriega, *Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, Tumer, Yolum, Sonenberg and Stone (eds.), May, 2–6, 2011, Taipei, Taiwan, pp. 1191-1192.

2. NORMATIVE MULTI-CONTEXT GRADED BDI ARCHITECTURE (N-BDI)

A logical multi-context system [3] is defined as a set of interconnected contexts. Each context has its own language and, typically, a modal logical system with axioms and inference rules. Contexts are connected through *bridge* inference rules whose premises and conclusions belong to different contexts. It is assumed that logical multi-context systems have computational implementations. The n-BDI architecture for *norm aware* agents that we propose (detailed in [2]) is formed by (see Figure 1):

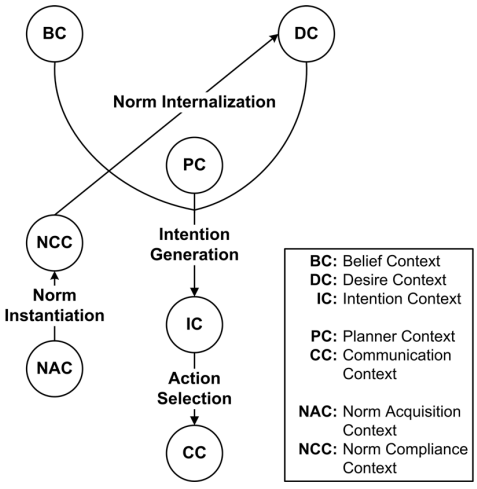

Figure 1: The n-BDI Architecture

1. **Mental contexts** that characterize beliefs (BC), intentions (IC), and desires (DC). Following [1], they are defined with propositional graded modal logics for

representing degrees of certainty, desirability, or intentionality of mental predicates.

2. We assume two **functional contexts** (also based on [1]) the Planner Context (PC), which allows agents to decide the set of actions that will be attempted according to their desires; and the Communication Context (CC), which communicates agents with their environment.

3. Finally, we include two **normative contexts** that allow agents to reason about an explicit representation of norms that are relevant for their actions [2]:

   - **Norm Acquisition Context** (NAC). It updates the set of norms that are in force at a given moment, i.e. the legislation the agent is subject to. Specifically, the NAC receives information from the environment (observed and communicated facts), determines if that information is a norm that regulates his own behaviour and updates, accordingly, his existing set of norms.
   - **Norm Compliance Context** (NCC). This is the component responsible for reasoning about the set of norms that hold at a specific moment. It determines those norms whose activation conditions are met. In this sense, the NAC contains all the abstract norms that are in force, whereas the NCC only contains those norm instances that are active in the current situation.

3. REASONING PROCESS IN THE N-BDI ARCHITECTURE

The n-BDI architecture described in [2] allows agents to have an explicit representation of norms. Thus, agents are capable of detecting the activation of norms and selecting those plans that comply with active norms. However, a norm-aware agent may decide whether to comply with a norm or not. In this work we propose a coherence-based mechanism to enable such an agent to make that decision. Namely, this paper proposes carrying out the reasoning process in the n-BDI architecture in three steps:

Step 1. Norm-based Expansion. This first step consists in extending the agent's theory of mental propositions with those norms that become active as well as those norms that become inactive. In other words, this step creates a state of mind where the agent is to fulfil all applicable norms. The norm-based expansion process is made up of two phases: (i) *NCC update*, i.e. when the activation conditions of a norm in the NAC hold, the abstract norm is instantiated and included in the NCC: likewise, when a termination condition is satisfied, the norm instance is removed from NCC; and (ii) *norm internalization*, where norms, currently in NCC, are propagated –through bridge rules– to the agent's mental and functional contexts. Notice that the updating of NCC is the agent's truthful understanding of the norms that are objectively applicable to him. The consequences of applicable norms are propagated to the agent's mental and functional contexts (*internalized*) every time NCC is updated because his actions are triggered by his prevalent state of mind. [1]

Step 2. Coherence-based Contraction. The internalization process just described may produce deontic conflicts within each context. In those cases, the agent needs to address those conflicts so that he may take action. Specifically, our proposal employs *coherence* as a criterion for determining which propositions (both mental and normative) must be removed to resolve those conflicts. In fact, we use coherence to face three different problems: (i) deliberating about the coherence of desires in view of applicable norms; (ii) determining degrees of coherence in states with normative conflicts; and (iii) in each context, choose a subset of maximal coherence to resolve normative conflicts. Actually, the coherence-based contraction algorithm takes into account the following: (i) the beliefs that sustain the activation of norms and other beliefs that explain or contradict them; (ii) the norm instances and the conflict relationships among them; and (iii) the evaluation of the main goals as well as other goals that potentially facilitate them.

Step 3. Decision Making. Finally, intentions are generated by considering plans that achieve those desires belonging to the coherence maximizing set. These intentions will determine the next action to be performed by the agent. For the key decision of norm compliance, we will profit from Joseph's proposal [4] to enable n-BDI agents to choose the propositions that maximize the coherence of the context. [2]

4. ACKNOWLEDGMENTS

This paper was partially funded by the Spanish government under grants CONSOLIDER-INGENIO 2010 CSD2007-00022, TIN2009-13839-C03-01, TIN2008-06701-C03-03, TIN2008-04446 and by the FPU grant AP-2007-01256 awarded to N. Criado. This research has also been partially funded by the Generalitat de Catalunya under the grant 2009-SGR-1434 and Valencian Prometeo project 2008/051.

5. REFERENCES

[1] A. Casali. *On Intentional and Social Agents with Graded Attitudes.* PhD thesis, Universitat de Girona, 2008.

[2] N. Criado, E. Argente, P. Noriega, and V. Botti. Towards a Normative BDI Architecture for Norm Compliance. In *COIN@MALLOW2010*, pages 65–81, 2010.

[3] C. Ghidini and F. Giunchiglia. Local models semantics, or contextual reasoning= locality+ compatibility. *Artificial intelligence*, 127(2):221–259, 2001.

[4] S. Joseph. *Coherence-Based Computational Agency.* PhD thesis, Universitat Autònoma de Barcelona, 2008.

[1] The "state of mind" is the union of the contents of all the contexts, in a norm-aware agent, these include normative elements. Up to now we have only considered the internalization of norms as goals; i.e., the NCC updates the DC with normative desires; these normative desires influence the agent's choice of the most suitable intended plan.

[2] In fact, [4] proposes a formalisation of the notion of coherence for multi-context graded BDI agents together with mechanisms for calculating the coherence of a set of graded mental attitudes.