A BDI Architecture for Normative Decision Making (Extended Abstract)

N. Criado, E. Argente, V. Botti
Grupo de Tecnología Informática - Inteligencia Artificial
Universidad Politécnica de Valencia (Spain)
{ncriado,eargente,vbotti}@dsic.upv.es

ABSTRACT

Norms, to become effective, must be recognised as norms by agents. These agents must be able to accept norms but maintaining their autonomy. In this paper, the multi-context BDI agent architecture has been extended with a recognition and a normative context in order to allow agents to acquire norms from their environment and consider norms in their decision making processes.

Categories and Subject Descriptors

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

General Terms

Theory

Keywords

Norms, Reasoning, BDI agents, Multi-Context systems

1. INTRODUCTION

Open MAS are characterised by a high uncertainty and limited trust in their performance. In this sense, norms have been proposed as mechanisms which guide agent behaviours and control the system performance. However, norms, to become effective, must be recognised as norms by agents. In addition, agents must be able to accept norms but maintaining their autonomy [2]. The norm acceptance problem consists of two related activities: the recognition of norms as such inside agent minds; and the consideration of these norms in the decision making process.

Usually, proposals on agent architectures which support normative reasoning do not consider norms as dynamic objects which may be acquired and recognised by agents. On the contrary, these proposals consider norms as static constraints that are hired on agent architectures. In this paper, the multi-context BDI agent architecture [1] has been extended with recognition and normative reasoning capabilities in order to allow agents to acquire norms dynamically and consider them in their decision making process.

Cite as: A BDI Architecture for Normative Decision Making (Extended Abstract), N. Criado, E. Argente and V. Botti, *Proc. of 9th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, van der Hoek, Kaminka, Lespérance, Luck and Sen (eds.), May, 10–14, 2010, Toronto, Canada, pp. 1383-1384

2. NORMATIVE BDI ARCHITECTURE

The norm acceptance problem consists of two main problems: norm recognition and norm decision [2]. In this sense, agents must be able to acquire new norms and deliberate about norm adoption autonomously. A multi-context BDI [1] agent is mainly formed by (Figure 1 grey units): mental units to characterize beliefs (BC), intentions (IC) and desires (DC); and functional units for planning (PC) and communication (CC). Taking as a reference this architecture, our proposal consists in extending it by adding two functional contexts (Figure 1, white units): the *Recognition Context* (RC), which is responsible for the norm identification process; and the *Normative Context* (NC), which allows agents to consider norms in their decision making processes.

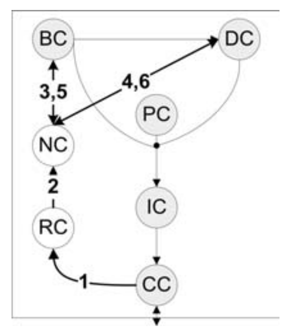

Figure 1: Multi-Context Normative BDI Architecture. The normative extensions are the white contexts and bold bridge rules.

2.1 Recognition Context (RC)

The norm decision process starts when the RC derives a new norm through analysing its environment. More concretely, the RC context receives the environmental facts, both observed and communicated, and identifies the set of norms which control the agent environment. This context is formed by expressions which are defined as $(RC\ \alpha, \delta_\alpha)$; where α is a first order formula which represents a norm and $\delta_\alpha \in [0, 1]$ is the certainty degree ascribed to the norm.

2.2 Normative Context (NC)

These recognised norms are translated into a set of inference rules which are included into the NC. The NC is responsible for deriving new beliefs and desires according to the current agent mental state and the inference rules which have been obtained from norms. These new beliefs and desires may cause intentions to be updated and, as a consequence, normative actions might be carried out. The NC is formed by expressions like $NC(\lceil\gamma\rceil)$; where γ is an inference rule $(\varphi \rightarrow \psi)$ which relates mental attitudes of an agent. $\lceil\gamma\rceil$ means that γ is embedded in the NC as a term.

2.3 Bridge Rules

Updating the RC. Agent observations and communications ($input(\beta)$) are included into the RC as a theory ($\lceil input(\beta)\rceil$) (see Figure 1 Rule 1):

$$CC : input(\beta) \Rightarrow RC : \lceil input(\beta)\rceil \tag{1}$$

Norm Transformation Rules. Inside the recognition unit new norms are acquired. Those abstract recognised norms ($RC\ \alpha, \delta_\alpha$) are transformed into terms belonging to the normative context ($NC(\lceil\gamma\rceil)$) (see Figure 1 Rule 2):

$$RC : (RC\ \alpha, \delta_\alpha) \Rightarrow NC : NC(\lceil\gamma\rceil) \tag{2}$$

As previously argued, each abstract norm is translated into an inference rule belonging to the NC. Next, bridge rules for translating each type of norm are described:

- *Constitutive Norm Transformation Rule*:

$$RC : (RC\ \langle count-as, A, E, \alpha \rightarrow \gamma\rangle, \delta_{nr}) \\ \Rightarrow NC : NC(\lceil\varphi \rightarrow \psi\rceil)$$

  where:

$$\varphi = (B\ A, \delta_A) \wedge (B\neg E, \delta_E) \wedge (\Phi\alpha, \delta_\alpha) \\ \psi = (\Phi\ \gamma, f(\delta_\alpha, f_{activation}(\delta_A, \delta_E, \delta_{nr})))$$

  A constitutive norm ($\langle count-as, A, E, \alpha \rightarrow \gamma\rangle$) represents that if the norm is active, according to its activation (A) and expiration (E) conditions, then the basic concept α can be redefined as the abstract concept γ. Thus, if an agent considers that the norm is currently active ($(B\ A, \delta_A) \wedge (B\neg E, \delta_E)$) and the basic fact α, affected by the constitutive norm, is an agent belief or desire ($(\Phi\alpha, \delta_\alpha)$) then a new belief or desire will be inferred corresponding to the new abstract fact ($\Phi\ \gamma$). $f_{activation}$ is a function that combines the belief degrees related to the norm conditions (δ_A and δ_E) and the certainty degree of the norm (δ_{nr}). The certainty degree, which is related to the norm activation, together with the certainty or desirability degree assigned to the basic fact (δ_α) are employed by the function f in order to assign a degree to γ.

- *Procedural Norm Transformation Rule*:

$$RC : (RC\ \langle D, A, E, C, S, R\rangle, \delta_{nr}) \Rightarrow \\ NC : NC(\lceil\varphi \rightarrow \psi\rceil)$$

  where:

$$\varphi = (B\ A, \delta_A) \wedge (B\neg E, \delta_E) \wedge (D^*C, \delta_C) \\ \wedge (D^- S, \delta_S) \wedge (D^+ R, \delta_R) \\ \psi = (D^*C, f(f_{adoption}(\delta_C, \delta_S, \delta_R), \\ f_{activation}(\delta_A, \delta_E, \delta_{nr})))$$

  A procedural norm ($\langle D, A, E, C, S, R\rangle$) determines a deontic control ($D \in \{Obliged, Forbidden\}$) over a condition C which is enforced through sanctions (S) and rewards (R). It is translated by this bridge rule into an inference rule which adds a new desire (D^*C) if the norm is active according to the current state. $* \in \{+, -\}$ is the sign ascribed to the new desire inferred from the norm. In case of obligation norms a positive desire of achieving the norm condition is inferred. On the contrary, a prohibition is transformed into an inference rule which asserts a negative desire if the norm is active. The degree of the new desire is determined by a function (f) that combines the certainty degrees assigned to norm activation ($f_{activation}$) and the desirability of norm compliance assigned by the norm adoption function ($f_{adoption}$). The $f_{adoption}$ takes as input parameters the desirability of the norm condition (δ_C), sanction (δ_S) and reward (δ_R).

Updating the NC. Agent desires and beliefs (γ) are included into the normative context ($\lceil\gamma\rceil$) in order to determine when a norm is active (Figure 1 Rules 3 and 4):

$$BC : \gamma \Rightarrow NC : NC(\lceil\gamma\rceil) \tag{3}$$

$$DC : \gamma \Rightarrow NC : NC(\lceil\gamma\rceil) \tag{4}$$

Updating the Mental Contexts. After performing the inference process for creating new beliefs ($\lceil(B\ \gamma, \delta)\rceil$) and desires ($\lceil(D^*\ \gamma, \delta)\rceil$) derived from norm application; the NC must update mental contexts (Figure 1 Rules 5 and 6):

$$NC : NC(\lceil(B\ \gamma, \delta)\rceil) \Rightarrow B : (B\ \gamma, \delta) \tag{5}$$

$$NC : NC(\lceil(D^*\ \gamma, \delta)\rceil) \Rightarrow D : (D^*\gamma, \delta) \tag{6}$$

3. CONLUSION

In this paper, the multi-context BDI agent architecture has been extended with a recognition and a normative contexts in order to allow agents to acquire these norms from their environment and consider them in their decision making process. The fact that mental attitudes of agents are quantified allows them to reason in open environments which are controlled by norms. In this sense, graded modalities allow agents to represent uncertain knowledge about the current state of the world. Moreover, graded intentions and desires enable agents to make decisions according to their satisfaction criterion. This is specially interesting when designing normative agents which behaviour can be affected by conflicting norms. Thus, the desirability degrees of desires and intentions allow agents to decide between norm violation or fulfilment according to their priorities.

4. ACKNOWLEDGMENTS

This work was partially supported by the Spanish government under grants CONSOLIDER-INGENIO 2010 CSD2007-00022, TIN2009-13839-C03-01 and TIN2008-04446 and by the FPU grant AP-2007-01256 awarded to N. Criado.

5. REFERENCES

[1] A. Casali, L. Godo, and C. Sierra. Graded BDI models for agent architectures. In *CLIMA V*, volume 3487 of *LNCS*, pages 126–143. Springer, 2004.

[2] R. Conte, C. Castelfranchi, and F. Dignum. Autonomous norm acceptance. *LNCS*, 1555:99–112, 1999.