Extending the BDI architecture with commitments

Dorian Gaertner [a,1], Pablo Noriega [a] and Carles Sierra [a]
[a] *Institut d'Investigació en Intelligència Artificial, IIIA - CSIC*

Abstract. In this paper, we describe a novel agent architecture for normative multi-agent systems which is based on multi-context systems. It models the three modalities of Rao and Georgeff's BDI agents as individual contexts and adds a fourth one for *commitments*. This new component is connected to all other mental attitudes via two sets of bridge rules, injecting formulae into it and modifying the BDI components after reasoning about commitments. As with other normative approaches the need for methods to deal with consistency is a key concern. We suggest three forms of dealing with the truth maintenance problem, all of which profit from the use of multi-context systems.

Keywords. BDI agents, normative multi-agent systems, logics for agents

1. Introduction

Autonomous agents are an important development towards the achievement of many of AI's promises. Among the many proposed agent architectures are Rao and Georgeff's well-known BDI agents [19] that model mental attitudes of an agent, concretely *beliefs* (representing the state of the environment), *desires* (representing the state of affairs the agent wants to bring about) and *intentions* (representing the currently selected goals). *Multi-context systems*, devised by Giunchiglia and Serafini [12] to structure knowledge into distinct theories, allow us to define complex systems with different formal components and the relationships between them. Parsons et al. [18] use these systems to model the three BDI modalities as individual components (*contexts* or *units*) with bridge rules to describe the dependencies between them. We propose to extend the BDI agent model with a fourth component that keeps track of the *commitments* an agent has adopted. We view a commitment as a triple consisting of the entity that commits, the entity that the commitment is directed at and the content of the commitment. These entities can be individual agents, groups of agents or institutions. In this paper, we will follow the approach taken by Parsons et al. and model agents as multi-context systems. We describe how the commitment component is connected to the other three contexts via instances of two basic bridge rule schemata and suggest approaches to handle arising inconsistencies.

In the next section we are going to formally define the use of the terms *Commitment* and *Norm* that we are employing in this paper. Subsequently, we summarise multi-context systems and explain how we extended them. We show how our architecture lends

[1]Correspondence to: Dorian Gaertner, IIIA, Spain. Tel.: +34 935809570; E-mail: dorian@iiia.csic.es.

itself to modelling *normative* MAS and propose a novel way to operationalise norms. Section 4 is concerned with truth maintenance and consistency issues. Finally, we contrast our architecture with existing ones, present our future work and conclude.

2. Commitments and Norms

Norms, normative agents and normative multi-agent systems have received a lot of attention in recent years. Lopéz y Lopéz et al. [15] proposed a formal model of these concepts using the Z specification language. García-Camino et al. [11] have analysed the concept of norms in a society of agents and how norms are implemented in an electronic institution. In [8], Dignum et al. extend the BDI architecture to handle norms. They are using PDL, a deontic logic, to formalise obligations from one agent to another. Norms, in their view, are obligations of a society or organisation. They explicitly state that a norm of a society is a conditional (p should be true when q is true). Finally, Cohen and Levesque in their paper 'Intention is choice with commitment' [6] talk about internal commitments as a precursor to the social commitments that we concern ourselves with.

We consider a *commitment* from one entity to be directed at another entity. With respect to these entities, one needs to distinguish between individual agents and groups of agents - or electronic institutions [1]. For example, an agent can be committed to an (electronic) institution to behave in a certain way. The institution on the other hand, may be committed to the agent to reward or punish him, depending on his behaviour. Note, that this is different from the case where one agent is responsible for norm enforcement. Commitments may also exist between agents or between different electronic institutions.

The content of a commitment can be a certain contract (e.g. an intention to deliver ten crates of apples once the agent believes to have been paid) between two agents or it can be a norm (e.g. you should not desire your neighbour's wife). In this paper, we will focus on the latter. The BNF description of our commitment language is hence (where WFF stands for a well-founded formula in an appropriate language, see e.g. [13]):

$$
\begin{aligned}
Commitment &::= Commit(S, S, \lceil C \rceil) \\
S &::= agent \mid institution \\
C &::= Contract \mid Norm \\
Contract &::= WFF \\
Norm &::= \varphi \rightarrow \psi \\
\varphi &::= ConjLiterals \\
ConjLiterals &::= MLiteral \mid MLiteral \wedge ConjLiterals \\
\psi &::= MLiteral \\
MLiteral &::= MentalAtom \mid \neg MentalAtom \\
MentalAtom &::= \mathtt{B}(term) \mid \mathtt{D}(term) \mid \mathtt{I}(term)
\end{aligned}
$$

We consider a *norm* to be a conditional, first-order logic formula that relates mental attitudes of an agent. All variables are implicitly universally quantified. In this paper, we are using beliefs ($\mathtt{B}$), desires ($\mathtt{D}$) and intentions ($\mathtt{I}$) to model an agent's mental state. For example,

$$\mathtt{B}(sex(X,S)) \wedge \mathtt{B}(sex(Y,S)) \rightarrow \neg\mathtt{I}(danceWith(X,Y))$$

is a norm which can be read as “for any two agents, if they have the same gender, they should not intend to dance together" (example taken from the social ballroom of Gaertner et al. [10]). The argument to the mental literals can be any term and the implication arrow of the norm can always informally be translated with the English word *should*.

Although, in the above formula the modalities B and I should have a subscript x indicating that we are talking about beliefs and intentions of agent x we drop these subscripts for readability whenever it is clear from the context which agent is referred to. Furthermore, we never need distinct subscripts in the same norm formula, since it does not make sense to say a belief of one agent causes an intention for another agent.

A norm, for us, is a social phenomenon, in that it applies to all agents in a given society or institution. Each agent is then committed to the institution to obey the norm. We therefore stipulate that if φ is a norm in an institution Π, the following must hold:

$$\forall \alpha : agent \in \Pi : inst \,.\, Commit(\alpha, \Pi, \lceil \varphi \rceil)$$

where $\lceil \varphi \rceil$ is the codification of a norm as a term in Gödel’s sense. Contrast this with the notion of a contract, which in most cases affects only two parties (or agents):

$$Contract(\alpha, \beta, \lceil \varphi \rceil) \rightarrow JointCommit(\{\alpha, \beta\}, \lceil \varphi \rceil)$$

where the notion of a joint commitment can be defined in arbitrarily complex ways. In what follows, we will mostly talk about agents who are committed to the institution they belong to. We therefore drop this information (i.e. the first two parameters) for brevity’s sake. Unless otherwise stated, a commitment to φ of the form: $Commit(\lceil \varphi \rceil)$ should be read as: $Commit(self, myInstitution, \lceil \varphi \rceil)$.

3. Multi-context Architecture: BDI+C

Multi-context systems (MCS) have first been proposed by Giunchiglia and Serafini in [12] and were subsequently used in a generic agent architecture by Noriega and Sierra in [16]. Individual theoretical components of an agent are modelled as separate *contexts* or units, each of which contains a set of statements in a language L_i together with the axioms A_i and inference rules Δ_i of a (modal) logic. A unit is hence a triple of the form:

$$unit_i = \langle L_i, A_i, \Delta_i \rangle$$

Not only can new statements be deduced in each context using the deduction machinery of the associated logic, but these contexts are also inter-related via *bridge rules*, that allow the deduction of a formula in one context based on the presence of certain formulae in other, linked contexts. An agent is then defined as a set of context indices I, a function that maps these indices to contexts, another function that maps these indices to theories (providing the initial set of formulae in each context) together with a set of bridge rules BR as follows:

$$Agent = \langle I,\ I \rightarrow \langle L_i, A_i, \Delta_i \rangle,\ I \rightarrow T_i,\ BR \rangle$$

The BDI+C agent architecture we are proposing in this paper extends Rao and Georgeff’s well-known BDI architecture with a fourth component which keeps track of the commitments of an agent. Below we describe a BDI+C agent as a multi-context system being inspired by the work of Parsons, Sierra and Jennings (see [18]). Each of the mental attitudes is modelled as an individual unit. Contexts for communication and planning (a functional context) are present in addition to the belief-, desire-, intention- and commitment-context but in this paper we will focus on the latter four.

For the belief context, we follow the standard literature (see for example, [17] and [19]) and choose the modal logic KD45 which is closed under implication, provides consistency, as well as positive and negative introspection. However, it does not have veridicality, which means that the agent's beliefs may be false. However, in such a situation the agent itself is not aware that its beliefs are false. Like Rao and Georgeff [19] we also choose the modal logic KD to model the desire and intention components.

For the commitment context, the logic consists of the axiom schema K, closure under implication, together with the consistency axiom D. This does allow for conflicting commitments, but prohibits to be committed to something and not be committed to that something at the same time. That means, that we do not allow both $Commit(\lceil\varphi\rceil)$ and $\neg Commit(\lceil\varphi\rceil)$ to be present at the same time. However, it is perfectly possible to have both $Commit(\lceil\varphi\rceil)$ and $Commit(not\lceil\varphi\rceil)$ in the commitment context . Due to the existence of schema K in this context, one can derive $Commit(\lceil\varphi\rceil\ and\ not\lceil\varphi\rceil)$ which is different from $Commit(false)$. The argument is just a term—we could assign any semantics to it. The beauty of MCS is that it allows us to embed one logic as terms into another logic (the logic of the component). Therefore, $Commit(and(\varphi, not\ \varphi))$ evolves to $Commit(false)$ only if we apply propositional logic to the language modelled in this context. The two introspection axioms do not apply, since it does not make sense to say that once an agent is committed to some cause, it is also committed to be committed to this cause; similarly for negative introspection. All units also have modus ponens, uniform substitution and all tautologies from propositional logic.

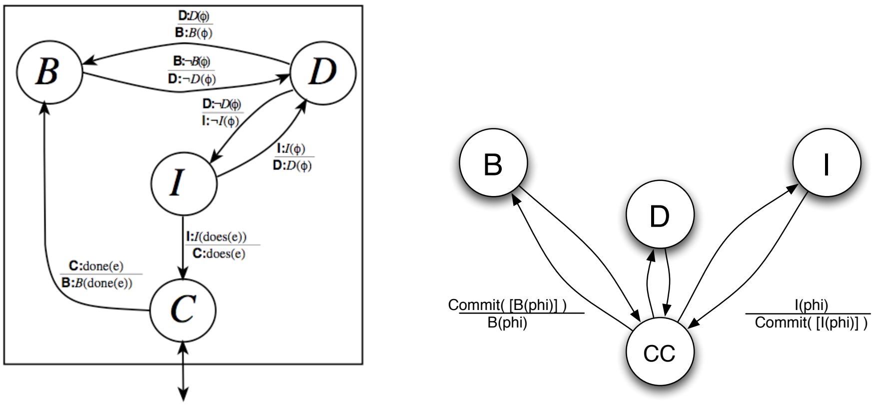
3.1. Bridge Rules

There are a number of relationships between contexts that are captured by so-called bridge rules. A bridge rule of the form (sometimes written on one line for convenience):

$$\frac{u_1 : \varphi,\ u_2 : \psi}{u_3 : \theta} \qquad \text{or} \qquad u_1\ :\ \varphi,\ u_2\ :\ \psi \rightarrow u_3\ :\ \theta$$

can be read as: if the formula φ can be deduced in context u_1 and ψ in u_2 then the formula θ is to be added to the theory of context u_3. It allows to relate formulae in one context to those in another. In [18] three different sets of bridge rules are described which model realistic, strongly realistic and weakly realistic-minded agents. Figure 1(a) shows the model of an strongly realistic agent. Note, that in these figures, the *C* represents the communication unit and *CC* the commitment unit. One of its bridge rules, for example, states that something that is not desired should also not be intended.

In addition to the information-propagating bridge rules in the figure, there are more complex rules related to awareness of intention and impulsiveness between the belief and intention units (see [17]). These are common to all strongly realistic agents. Finally, there are domain specific rules, which link the contexts to the communication unit and control the impact of interaction with the environment on the mental state of an agent. An example of this is the bridge rule that stipulates that everything that is communicated to an agent to be done is believed to be done.



(a) (b)

Figure 1. (a) Multi-context description of a strongly realist BDI agent taken from [18]. Note, that *C* stands for a standard *communication* context as the authors did not concern themselves with commitments. (b) Commitment overlay for normative agents. In this figure, square brackets represent Goedel codification.

We are proposing to add an extra layer of bridge rules to existing BDI multi-context agents that controls the content of the mental contexts via norms. Remember, we earlier stated, that an adopted norm becomes a conditional commitment. The default *normative* personality of an agent is expressed as follows: an agent commits to believe everything it believes, commits to desire everything it desires and commits to intend everything it intends; and an agent beliefs what it is committed to believing, desires what it is committed to desiring and intends what it is committed to intending. This is modelled via two sets of bridge rules where Φ stands for any of `B`, `D` or `I`:

$$\frac{\Phi : \Phi(\varphi)}{CC : Commit(\lceil \Phi(\varphi) \rceil)} \; (*) \qquad\qquad \frac{CC : Commit(\lceil \Phi(\varphi) \rceil)}{\Phi : \Phi(\varphi)} \; (**)$$

Two examples of this can be seen in Figure 1(b) that depicts the normative layer of bridge rules we propose. Formulae (restricted to mental literals) from any of the three standard contexts are injected into the commitment context via a bridge rule of the form (*), where they encounter norms (first-order logic implications). Since the commitment context is closed under implication, the deduction machinery inside this context can be thought of as *applying* the norms. The resulting formulae of the local reasoning in the commitment unit are then injected back into the appropriate context via a bridge rule of the form (**).

The six arcs in the figure represent the *default* normative personality of an agent. It is perfectly reasonable to imagine agents with different attitudes towards norms. A rebellious agent for example, may not desire or intend everything that it is committed to desiring or intending. Modelling agent types can therefore proceed on two levels. At the standard level between the belief-, desire- and intention context, personality traits like strong realism can be modelled, whereas character traits related to norms and norms adoption can be mimicked by modifying the overlay net of bridge rules involving the commitment context.

This proposed architecture is operationally speaking very simple. The complexity of norm execution is dealt with in the commitment context, whose logic is easily modifiable. Our modular, layered approach is a natural, clean extension that provides BDI agents with a new capability, namely norm compliance.

4. Truth Maintenance Problem

Adopting a norm and hence adding a commitment is in some way like opening a channel, linking different parts in the agent's 'brain'. For example, a commitment of the form $Commit(\ulcorner \mathtt{B}(\varphi) \rightarrow \mathtt{I}(\psi) \urcorner)$ causes $\mathtt{I}(\psi)$ to be deduced in the intention context if $\mathtt{B}(\varphi)$ is deducible in the belief context. The reasoning is as follows:

- a bridge rule from the normative layer (see (*) in Section 3) adds $Commit(\ulcorner \mathtt{B}(\varphi) \urcorner)$ to the commitment context since $\mathtt{B}(\varphi)$ is deducible in the belief context
- the adopted norm together with an instance of schema K allow us to deduce $Commit(\ulcorner \mathtt{I}(\psi) \urcorner)$
- another normative bridge rule (see (**)) injects $\mathtt{I}(\psi)$ into the intention context

One can therefore think of this as having a bridge rule linking the belief and the intention context, which is only activated, once $Commit(\ulcorner \mathtt{B}(\varphi) \rightarrow \mathtt{I}(\psi) \urcorner)$ is present in the commitment context. What happens however, if $\mathtt{B}(\varphi)$ is removed from the belief context? Should one also remove $\mathtt{I}(\psi)$ from the intention context? What impact does the revocation of the commitment have? In any case, one has to ensure the consistency of all the mental attitudes (since their respective logics contain schema D). Generally, *adopting a norm* has extensive ramifications with respect to consistency. This dilemma is known as the *truth maintenance (TM) problem*. Artificial Intelligence has seen many different approaches to the TM problem. For our purposes, the most promising ones are:

4.1. Standard truth maintenance systems

Once a bridge rule has fired and tries to inject a formula into a context, it is the responsibility of the context to maintain consistency. In the simplest case, it checks, whether the formula to be inserted is inconsistent with the existing theory of the context and if it finds this to be the case, rejects the proposed injection. This is a very simplistic approach that only allows monotonic updates. More sophisticated truth maintenance systems can handle non-monotonic updates or *belief revision*. A formula which contradicts the existing theory in a context can still be inserted, but some machinery must then revise the theory to make it consistent again (by removing some of the causes of the contradiction). Two main approaches whose use we are investigating currently are *justification-based* truth maintenance systems like the one proposed by Doyle [9] and *assumption-based* truth maintenance systems following the work by deKleer [7].

4.2. Argumentation

A traditional way of resolving conflicts is to consider the arguments in favour and against a decision and choose those that are more convincing. The area of *argumentation* studies this process and gives tools, mostly based on logical approaches, to automate this decision process (see for example the work detailed in [3] and [18]). In these works the decision is made considering that arguments are proofs in a logical formalism and that the proofs *attack* one another by deducing opposite literals or *rebut* one another by deducing the opposite of a literal used to support a proof.

In our case we need to have a notion of argument that bases the *attack* not *only* on some logical relationships between the proofs used to support two opposite literals, but also the fact that some of the proofs are based on the application of norms. Therefore, beyond the logical attack one has to consider the *strength* of the argument in how supported it is by the norms of the institution. In that sense, we suggest to include in the proof the set of norms applied to generate a commitment and use a measure over them when the content of the commitment challenges a pre-existing intention, belief or desire. This measure can be based on specific reasoning with respect to the willingness that the agent has to respect a given norm, or the *degree of adoption* of the norm by the agent.

4.3. Decision Theory and Graded Mental Attitudes

Decision theory on the other hand is based on *utilities*. When faced with two conflicting intentions, one dictated by a norm and the opposite dictated by the agent's desires, it may decide to violate the norm, if this violation and the fulfilment of its desire are more *satisfying* then conforming to the norm. In order to decide what is more satisfying, we propose to use graded mental attitudes similar to the work done by Casali et al. [4].

In their work, the atomic formulae in contexts are no longer of the form $\mathsf{B}(term)$ but instead are enriched with a weight ϵ to give $\mathsf{B}(term, \epsilon)$. This weight represents the degree of belief. Similarly, for desires, it represents the degree of desire allowing us to attach priorities to certain formulae. In the case of intentions, the weight can be used to model the cost/benefit trade-off of the currently intended action. Finally, a weight on a commitment indicates the degree of adoption. Using these graded modalities, one can compute the utility of each of the conflicting atoms and act accordingly.

5. Related Work

We have referred to related work throughout the paper, in particular in the first half. In this section, we aim to contrast our proposal with two particular lines of work, namely a modified BDI interpreter by Dignum et al. [8] and the BOID architecture by Broersen et al. [2]. Dignum and his colleagues add one step to the main loop of the BDI interpreter in which selected events are augmented with deontic events by repeatedly applying the introspective norms and obligations [8]. They distinguish between norms (that hold for a society) and obligations (that hold between two agents). They rank obligations based on the punishments associated with their violation and norms based on their social benefit. Our view of commitments is broader in that we allow the committed entities and the subjects of the commitment to be agents, groups of agents or entire societies. The architecture we propose is more flexible, too, since each component has its own logic and the relationships between components can be varied dynamically.

The BOID architecture by Broersen and his colleagues has many similarities with our work. It contains four components (B, O, I and D) where the O component stands for obligations (as opposed to commitments in our case) and the other components have the usual meaning. They suggest feedback loops that feed the output of every component back to the belief component for reconsideration. The order in which components are chosen for rule selection, determines the kind of character the agent possesses. For example, if obligations are considered before desires, the agent is deemed to be social. One drawback is, that they only consider orders in which the belief component overrules any other modality [2]. Furthermore, these orders are fixed for each agent. Using our agent architecture, agent types can be modified dynamically. Also, the relationship between mental attitudes can be controlled at a finer level of granularity (e.g. domain-specific rules connecting multiple contexts rather than the strict ordering of components required in the BOID architecture).

6. Future Work

We are currently working on implementing the BDI+C agent architecture using QuP++ [5] an object-oriented extension of QuProlog. The advantage of this particular Prolog variant is its multi-threadedness and support for reasoning. We are implementing every context as an individual thread and use separate threads for bridge rules to synchronise between the contexts. Another line of research is concerned with generalising both the architecture and the implementation to handle graded mental attitudes. Casali et al. [4] have formalised the notion of uncertainty for the BDI model and we believe it can be employed in the BDI+C model in order to tackle the truth maintenance problem. Furthermore, it will allow us to represent the character or type of an agent more closely. We even envisage the ability to express the *mood* of an agent via dynamically changing the degree to which it believes, desires, intends and sticks to its commitments.

Furthermore, our interest lies in investigating temporal aspects of norms and norm adoption. In [20], Sabater et al. extend the syntax of bridge rules by introducing the notions of consumption and time-outs. We intend to make use of these extensions in order to allow for more expressiveness in the formulation of normative commitments.

Lopéz y Lopéz, in her doctoral thesis [14], describes different strategies for norm adoption ranging from fearful, rebellious and greedy character traits to reciprocation and imitation of other agents. All her strategies are based on potential rewards or punishments. Broersen et al. in [2] define agent characters based on the fixed order of the belief-, obligation-, intention- and desire-component (though they do not use multi-contexts, one can think of their components as such). They also give names such as 'super-selfish' to some of these orderings. Using the extended bridge rule layer of our architecture combined with graded versions of the mental attitudes, we can define different agent characters more formally and on a much finer level of granularity. The notions of release from commitment and norm evolution are also very interesting in this context. We intend to stretch the applicability of our proposed agent architecture to find out its limitations and possibly expand it.

7. Conclusions

In this paper we have outlined a conservative extension of BDI agent architectures to grasp the notion of commitments. We have proposed how to make these extensions operational in terms of multi-context logics and illustrated them with an example of dance negotiations following the etiquette conventions of a ballroom (more details about social norms and etiquette can be found in [10]). We found that our proposed extension of a BDI architecture—to incorporate commitment—is easy to formalise and make operational and has the following features:

1. It may be readily added on top of a given BDI model by simply including a new context and bridge rule schemata linking it to each of the other modalities.
2. Although we have proposed a schema that is uniform for all modalities, it is easy to fine-tune any given formalisation of the features of the commitment unit and the underlying BDI architecture in order to capture alternative formalisations, shades of meaning or the character or personality of an agent.
3. Our BDI+C model appears to be general enough to explore with it the complex aspects of legal consequence; especially in its concrete aspects of individual norm compliance with respect to the attitude of an agent towards authority, utility, selfishness and other features that have been addressed by the MAS community.
4. The notion of norms as an initial theory for the commitment context and the commitment-dependent bridge rules provide convenient ways to study paranormative aspects like norm adoption, compliance, blame assignment, violation, reparation or hierarchical normative sources. Likewise, the notion of contract could be modelled as joint commitments and added to the commitment context.
5. In a similar fashion, we have only pointed out a straightforward translation of norms as commitments between individuals and an institution, although it should be evident that other notions of authority (hierarchies of norms, issuers of norms, contingent applicabilities of norms) may be modelled along the same lines.
6. The evolution of the belief-, desire-, intention- and commitment theories as interaction proceeds and associated consistency issues may be addressed with the type of tools that have been applied to other dynamic theories, although in this paper we only hinted at three mechanisms: standard truth-maintenance systems, graded versions of the modalities and argumentation.

Acknowledgements

The authors would like to thank the Spanish Ministry of Education and Science (MEC) for support through the Web-i-2 project (TIC-2003-08763-C02-00). The first author is also grateful for a student grant from the Spanish Scientific Research Council through the Web-i(2) project (CSIC PI 2004-5 0E 133).

References

[1] Josep Ll. Arcos, Marc Esteva, Pablo Noriega, Juan Antonio Rodríguez, and Carles Sierra. Environment engineering for multiagent systems. *Journal on Engineering Applications of Artificial Intelligence*, 18(2):191–204, 2005.

[2] Jan Broersen, Mehdi Dastani, Joris Hulstijn, Zisheng Huang, and Leendert van der Torre. The BOID architecture: conflicts between beliefs, obligations, intentions and desires. In *AGENTS '01: Proceedings of the fifth international conference on Autonomous agents*, pages 9–16, New York, NY, USA, 2001. ACM Press.

[3] Marcela Capobianco, Carlos I. Chesñevar, and Guillermo Simari. Argumentation and the dynamics of warranted beliefs in changing environments. *Intl. Journal on Autonomous Agents and Multiagent Systems (JAAMAS)*, 11:127–151, September 2005.

[4] Ana Casali, Lluis Godo, and Carles Sierra. Graded BDI models for agent architectures. In *5th International Workshop on Computational Logic in Multi-Agent Systems (CLIMA V)*, pages 126–143, Lisbon, Portugal, 2004.

[5] Keith L. Clark and Peter J Robinson. *Computational logic: logic programming and beyond: essays in honour of Robert A. Kowalski, Part I*, chapter Agents as multi-threaded logical objects, pages 33–65. Springer, 2002.

[6] Philip R. Cohen and Hector J. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42(2-3):213–261, 1990.

[7] Johan de Kleer. An assumption-based TMS. *Artificial Intelligence*, 28(2):127–162, 1986.

[8] F. Dignum, D. Morley, E. Sonenberg, and L. Cavendon. Towards socially sophisticated BDI agents. In E. H. Durfee, editor, *ICMAS '00: Proceedings of the Fourth International Conference on MultiAgent Systems (ICMAS-2000)*, pages 111–118, Washington, DC, USA, 2000. IEEE Computer Society.

[9] Jon Doyle. A truth maintenance system. *Artificial Intelligence*, 12(3):231–272, 1979.

[10] Dorian Gaertner, Keith L. Clark, and Marek J. Sergot. Ballroom etiquette: a case study for norm-governed multi-agent systems. In *Proceedings of the First International Workshop on Coordination, Organization, Institutions and Norms (COIN) 2006 at AAMAS, Hakodate, Japan*, 2006.

[11] Andrés García-Camino, Pablo Noriega, and Juan Antonio Rodríguez-Aguilar. Implementing norms in electronic institutions. In *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems (AAMAS 2005)*, pages 667–673, New York, NY, USA, 2005. ACM Press.

[12] Fausto Giunchiglia and Luciano Serafini. Multilanguage hierarchical logics or: How we can do without modal logics. *Artificial Intelligence*, 65(1):29–70, 1994.

[13] John Knottenbelt and Keith Clark. Contract related agents. In *Sixth International Workshop on Computational Logic in Multi-Agent Systems (CLIMA VI)*, 2005.

[14] Fabiola Lopéz y Lopéz. *Social Powers and Norms: Impact on Agent Behaviour*. PhD thesis, University of Southampton, 2003.

[15] Fabiola Lopéz y Lopéz, Michael Luck, and Mark d'Inverno. A normative framework for agent-based systems. In *Proceedings of the 1st International Symposium on Normative Multiagent Systems*, 2005.

[16] Pablo Noriega and Carles Sierra. Towards layered dialogical agents. In *ECAI '96: Proceedings of the Workshop on Intelligent Agents III, Agent Theories, Architectures, and Languages*, pages 173–188, London, UK, 1997. Springer-Verlag.

[17] Simon Parsons, Nicholas R. Jennings, Jordi Sabater, and Carles Sierra. Agent specification using multi-context systems. In *Selected papers from the UKMAS Workshop on Foundations and Applications of Multi-Agent Systems*, pages 205–226, London, UK, 2002. Springer-Verlag.

[18] Simon Parsons, Carles Sierra, and Nick Jennings. Agents that reason and negotiate by arguing. *Journal of Logic and Computation*, 8(3):261–292, 1998.

[19] Anand S. Rao and Michael P. Georgeff. BDI agents: From theory to practice. In *Proceedings of the First International Conference on Multiagent Systems (ICMAS)*, pages 312–319, San Francisco, California, USA, 1995.

[20] Jordi Sabater, Carles Sierra, Simon Parsons, and Nicholas R. Jennings. Engineering executable agents using multi-context systems. *J. Log. Comput.*, 12(3):413–442, 2002.