

Working Towards a BDI-Agent based on Personality Traits to Improve Normative Conflicts Solution

Paulo Henrique Cardoso Alves, Marx Leles Viana and Carlos José Pereira de Lucena

Laboratory of Software Engineering (LES) - Department of Computer Science
Pontifical Catholic University of Rio de Janeiro – PUC – Rio, Rio de Janeiro, RJ - Brazil
{palves, mleles, lucena}@inf.puc-rio.br

Abstract— Norms exist to avoid and solve conflicts, make agreements, reduce complexity, and in general to achieve a desirable social order. However, norms eventually can be conflicting — for example, when there is a norm that prohibits an agent to perform a particular action and another norm that obligates the same agent to perform the same action at the same period of time. The agent’ decision about which norms to fulfill can be defined based on rewards, punishments and agent goals. Sometimes, this balance will not be enough to allow the agent to make the best decision. In this context, this paper introduces an approach that considers the agent’s personality traits in order to improve the solving process of normative conflicts.

Keywords. *Solving Normative Conflicts, Normative Agents, Multi-Agent Systems.*

I. INTRODUCTION

Multi-agent Systems (MASs) are societies in which these heterogeneous and individually designed entities (agents) work to accomplish common or independent goals [1]. In order to deal with autonomy and diversity of interests among the different members, such systems provide a set of norms, which are mechanisms used to restrict the behavior of agents by defining what actions the agents are obligated, permitted, or prohibited to encourage the fulfillment of the norm through rewards definition and discouragement of norm violation by pointing out the punishments [2].

Norms must be complied with by a set of agents and include normative goals that must be satisfied by the addressees. In addition, norms are not always applicable, and their activation depends on the background in which agents are situated. In some cases, norms suggest the existence of a set of sanctions to be imposed when agents fulfill, or violate, the normative goal.

The decision-making process about which norms will be fulfilled or violated might be defined based on the agent’s goals, rewards and punishment analysis [1]. Since an agent’s priority is the satisfaction of its own goals, before complying with norms, the agent must evaluate their positive and negative effects on its goals [3] without hurting the agent’s autonomy. Both rewards and punishment are the means for the agents to know what might happen independently of the agent’s decision to comply, or not, with the norms. However, norms sometimes may conflict or be inconsistent with one another [4]. For instance, different norms can, at the same time, prohibit and obligate a state that the agent wants to fulfill and the simple balance between goals, rewards and punishment might not be enough for the agent to make the best decision.

The abstract normative agent architecture developed by [3], has four main steps: (i) agent perception, (ii) norm adoption, (iii) norm

deliberation, and (iv) norm compliance. Within the norm deliberation step, conflicting norms are verified and a set of these norms is added to the norm compliance set.

We changed the internal process of the norm deliberation step to deal with conflicting norms by adding the agent’s personality traits. These characteristics will help the software agents to make a better decision involving personality traits - for example, sense of duty and spiritual endeavor. We will present a user scenario that shows how the agents deal with normative conflicts when personality traits are considered. This will illustrate the new deliberation process proposed in this paper.

In this context, we present an approach to build emotional BDI-agents, which considers also others agent’s personality traits [5] and emotions [6] to improve the decision-making process in the solution of normative conflicts. This approach aims at providing new resources for the agent to deal with conflicting norms supported by personality traits. As such, more human characteristics can be considered to improve the deliberation process. We built a software framework based on this approach, which provides a set of hot-spots and frozen-spots that enables the implementation of emotional normative functions. By using these new functions, it is possible to build emotional agents that: (i) use personality traits to improve the solution among normative conflicts, (ii) implement the agent’s behavior similar to the human’s behavior, and (iii) evaluate the effects on its desires with respect to the fulfillment or violation of a norm.

The reminder of this paper is organized as follows: Section II focuses on the background, while Section III discusses related work. Section IV presents the emotional BDI approach to solve normative conflicts. Section V describes a case study applying the emotional approach. Finally, Section VI presents our conclusions and future work.

II. BACKGROUND

A. Norms

Norms are designed to regulate the behavior of the agent, and therefore, a norm definition should include the address of the agent being regulated [7]. However, norms are different from laws, and they cannot force agents to comply with them. Agents are autonomous entities, so norms may only suggest and present the expected behavior.

In this work, we used the norm representation described in [9], which is composed by the representation of the element *norm* – it contains many different properties: (i) Addressee, (ii) Activation, (iii) Expiration, (iv) Rewards, (v) Punishments, (vi) Deontic

Concept, and (vii) State. For example, the property *Addressee* is used to specify the agents or roles responsible for fulfilling the norm.

In order to better understand the definition of norms and their representation, a user scenario was developed.

B. Conflicting Norms

Norms eventually may conflict, i.e., an action may be simultaneously prohibited and permitted, or it may be inconsistent, i.e., when an action is simultaneously prohibited and obliged [4]. These conflicts and inconsistencies may be caused by a norm that prohibits an agent to perform a particular action while another norm requires the same agent to perform the same action at the same time. For example, Fig. 1 presents a scenario of conflicting norms - when a norm defines that the buyer agent cannot give back the product bought and at the same time another norm defines that the buyer agent can return the product bought before opening it.

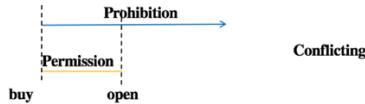


Figure 1. Conflict - Prohibition and Permission.

In short, conflicts may occur in different cases and situations, and dealing with them is extremely necessary to make the best decision.

C. BDI Architecture

The BDI (*Belief-Desire-Intention*) model was proposed by [10] as a philosophical theory of practical reasoning, representing, respectively, the information, motivational and deliberative states of the agent. There are two main steps: (i) apply a filter to make a set of goals that the agent has to commit to base on his beliefs, and (ii) find a way to know how the desires produced can be fulfilled based on the available agent's resources [12].

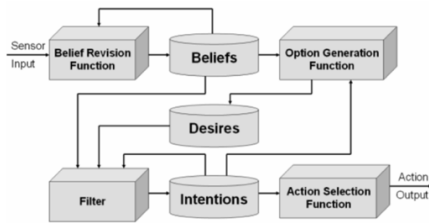


Figure 3. Generic BDI architecture [11].

The BDI model is composed by three mental states: (i) beliefs, which represent the environment factors that are updated after each action perceived — these beliefs represent the world knowledge; (ii) desires, which have information about the goals to be fulfilled — they represent the agent's motivational state, and (iii) intentions, which represent the action plan chosen. Fig. 3 shows these three mental states centralized and their interaction.

BDI architecture starts with a *Belief Revision Function* that makes a new belief set based on the agent's perception. Next, the *Option Generation Function* sets the agent's available options and desires, based on its own environment beliefs and intentions. The next function is a *Filter* that sets the agent's intentions based on its own beliefs, desires and intentions. Finally, the *Action Selection Function* sets the actions to be executed based on the current intentions.

Most BDI systems are inspired by the Rao e Georgeff [13] model. The authors presented an abstract BDI interpreter. This interpreter works with beliefs, goals and agent plans. As such, the goals are a set of concrete desires that may be evaluated altogether, avoiding a complex goal deliberation step. The interpreter's main activity is the means-end process achieved by plan selection and plan execution given a goal or event.

III. RELATED WORK

This section describes some related work: (i) the solution for normative conflicts [1], [8], [15], [16]; (ii) architecture designs considering the agent's emotional state [14], and (iii) the agent's personality [5].

The authors in [5] built a decision process to work as part of the story-telling systems wherein narrative plots emerge from the acting characters behavior and personality traits. The process evaluates goals and plans to examine the plan commitment issue. The drives, attitudes and emotions play a major role in the process. However, the personality traits were not applied on MASs, which creates an opportunity to improve the agent's decision-making process to deal with normative conflicts.

Some approaches [1], [8], [15], [16] have been proposed in the literature to develop the agent that evaluates the effects of solving normative conflicts. For instance, the n-BDI architecture defined by Criado *et al.* [15] presents a model for building environments governed by norms. Basically, the architecture selects objectives to be performed based on the priority associated with each objective. An objective's priority is determined by the priority of the norms governing a specific objective. However, it is not clear in this approach how the properties of a norm can be evaluated. In addition, the approach does not support a strategy and neither consider the agent's personality traits to deal with conflicts between norms.

Lopes *et al.* [8] defined a set of strategies that can be adopted by agents to deal with norms as follows: *Pressured*, *Opportunistic* and *Selfish*. Although this work provides some mechanisms for the agents to collect norms, the authors provide a framework that can be extended to create simulations of normative multi-agent systems by including new strategies. In addition, this work can neither extend mechanisms to collect information during the simulations nor can extend mechanisms to generate norms and agent goals. Furthermore, the agent cannot detect and overcome normative conflicts.

Finally, Viana *et al.* [1] presents a modeling language and an architecture to build adaptive normative agents. The authors propose an approach to design and implement agents that are capable to adapt in order to deal with norms, detecting and overcoming normative conflicts. However, this research just measures norms contributions based on: (i) norm rewards and punishments; (ii) norm activation and expiration; (iii) deontic concept, and (iii) agent goals. As such, the agent can decide to fulfill or violate a norm. One item that was not broached by the authors is that they did not implement personality traits in their architecture to improve and overcome normative conflicts.

As none of this related work deals with norms conflicts using personality traits, this was the gap that we based on to propose our work. We aim at providing a better way to balance goals, rewards, punishment and personality traits to solve normative conflicts. To evaluate the norm contribution, we first use rewards and punishment values. With these values, we then continue to evaluate the norm contribution, now adding personality traits.

IV. EMOTIONAL BDI AGENTS: AN APPROACH

This section describes the main concepts required to understand the approach based on emotional BDI-agents used to improve the solution of normative conflicts. **Error! Reference source not found.**

A. The Architecture

The emotional BDI-agents that can solve the normative conflicts approach were inspired on the concepts presented in the *background* and the related work section. We added both BDI features and personality traits in the normative deliberation process, mainly in relation to conflicts resolution. The architecture foundation was based on the abstract normative agent architecture developed in [3]. Fig. 4 presents our emotional BDI-agent architecture to solve normative conflicts.

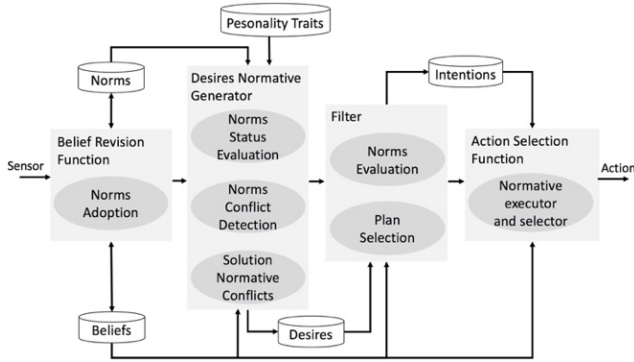


Figure 4. Internal architecture of the BDI-Agent based on Personality Traits to Improve Normative Conflicts Solution.

The most significant change was adding to the deliberation process a reasoning step that involves the BDI architecture and the personality traits approach. Both strategies work in a complementary way to make agents behavior more human, considering factors that were not used in the norms deliberation process in previous work. All of these changes refer only to the internal agent process. The decision-making process proposed has four steps, which is described below.

The first step involves the agent’s perception in the *Belief Revision Function*, where the agent perceives the norms in the environment addressed to it by means of sensors. Then, the agent inserts into the *Norms* set the norms that it wants to fulfill by using the *Norms Adoption* function. After that, the agent updates its beliefs, taking into account these new norms.

The second step is the *Desire Normative Generator*, which is composed by three processes: (i) *Norm Status Evaluation* function, where the agent verifies which norms are activated or deactivated; (ii) *Norms Conflict Detection* function, where the agent verifies what the normative conflicts are, and (iii) *Solution Normative Conflicts* function, where the agent evaluates the norms contribution and solves the normative conflicts, also considering its personality traits. There are some personality traits considered: (i) Drives, such as Sense of duty and Spiritual endeavor, (ii) Attitudes, such as Careful and Adaptable, and (iii) Emotions, such as Anger and Fear as presented in [5]. As a result, a set of non-conflicting norms are exported to the next step. These norms are the agent’s *Desires*.

The third step is the *Normative Filter*, which is composed by two processes: (i) *Norms Evaluation* function, where the agent evaluates the *Desires* set and it decides which norms will be

fulfilled, and (ii) *Plan Selection* function, where the agent’s best plans will be chosen in the *Intentions* set.

Finally, the fourth step is the *Action Selection Function*, which is composed by the *Normative executor and selector* function. This function receives the *Norms* set, which are the norms that the agent intends to fulfill. Last but not least, all of these steps help in the improvement of the normative conflict solving process, considering personality traits inserts into the BDI reasoning process.

B. The Framework

Inspired by the JSAN architecture [9], which uses different norms strategies to deal with norm, taking into account the different agent’s social levels, as in [8]. We built a new approach by introducing personality traits aiming to improve the solution of the normative conflict.

The solving process of normative conflicts starts with the calculation of the norm’s normative contribution, wherein for each norm the agent evaluates its rewards and punishments compared with the others norms addressed to it. Furthermore, we added a new step to improve this process, also taking into consideration the agent’s goals and its personality traits. This new step consists in the evaluation of which normative goals can be fulfilled according to the agent’s goals and its personality traits. The agent will verify which goals can be fulfilled based on its personality traits, so the agent uses its set of goals and analyze each conflicting norm, adding to the normative contribution an integer value to represent the compatibility between the agent’s goals and the normative goals. The compatibility is defined by the evaluation of which of the agent’s goals can be executed if a norm is fulfilled. As a result, some conflicting norms may have changed its normative contribution based on the use of the agent’s personality traits. For instance, imagine the norm that obliges the agent to cross a damaged bridge. If the agent is careful (careful meaning the agent’s personality trait) its normative contribution will be decreased because the agent does not have the intent to cross a damaged bridge — it is dangerous.

V. USER SCENARIO: GO HOME

As proof of concept, the user scenario “go home” will choose whether the agent goes home by bicycle, or by bus. The norms in this scenario are: (i) *Norm 1* prohibits the employee agent to go home by bicycle, and (ii) *Norm 2* obligates the employee agent to go home by bicycle. There are the *Norm 1* properties: (i) Name: Come back by Bus, (ii) Addressee: Employee, (iii) Deontic Concept: Prohibition, (iv) Reward: No health decrease, (v) Punishment: Be wet, (vi) Activation: It is raining, and (vii) Deactivation: It is sunny. *Norm 2* was defined by the following properties: (i) Name: Come back by Bicycle, (ii) Addressee: Employee, (iii) Deontic Concept: Obligation, (iv) Reward: Increase physical conditioning, (v) Punishment: Spend money with bus tickets, (vi) Activation: After work, and (vii) Deactivation: Be sick. Planning to go home, the employee agent checks the weather; if it is raining, it can go home by bus and as a consequence, it will violate *Norm 2*. However, if it is raining, but the employee agent has personality traits that induce its behavior to go home by bicycle, as a consequence, it will violate *Norm 1*. That is when the agent’s internal process detects and tries to overcome the normative conflict between *Norm 1* and *Norm 2*.

Fig. 6 shows the normative conflict between *Norm 1* and *Norm 2* and our aim is to present improvement in the deliberation process to choose the norm that will be fulfilled, considering some characteristics, such as: (i) the rewards of the norm that will be fulfilled; (ii) the punishment of the norm that will be violated; (iii)

the agent's goals, and (iv) the personality traits — for instance, if the agent's goal is to increase physical conditioning, it will have adventurous spirit as a personality trait.



Figure 6. Go Home conflict area.

All of these attributes were mapped to integers values in our architecture to make possible the decision process to choose between *Norm 1* and *Norm 2*. For comparison purposes, three different personality traits scenarios were developed for the employee agent: (i) adventurous spirit — *high* weight; (ii) adventurous spirit — *low* weight, and (iii) no personality trait.

We used the architecture proposed in this paper (see Fig. 4) in the first and second scenarios. For the three scenarios, we considered that the employee agent starts the *Norm Adoption* process to verify which norms are addressed to it. As a result, the employee agent perceives two norms. Note that these two norms are conflicting (see Fig. 6): both are active at the same time, and their deontic concept are opposite — obligation and prohibition. To choose which norms will be better fulfilled, the agent considers the normative contributions and its personality traits, except in the third scenario, where no personality trait is considered.

In the first scenario, we consider the employee agent with adventurous spirit — *high* weight to choose the norm that will be fulfilled in the conflict resolution process, also taking into account the norms punishments and rewards. The conflict resolution process measures, first of all, the norms rewards and punishments of each one of the conflicting norms, i.e., the agent verifies which goals can be executed, considering each one of the conflicting norms to be fulfilled. In sequence, the agent selects which norms will be fulfilled based on the agent's *Pressured* strategy. As a result, the employee agent will go home by bicycle. In the second scenario, we consider the employee agent with adventurous spirit — *low* weight. As a result, the employee agent decides that going home by bus will give it more benefits. This is because the agent has not enough motivation to fulfill its desires and as a consequence, it receives the punishments for not fulfilling the other norm.

Finally, in the third scenario, we consider the employee agent without personality traits, i.e., the agent always had the same behavior and considered only its own goals. We therefore assume that the BDI architecture with personality traits can change the agent's behavior, thus helping to improve the solution for the normative conflicts.

VI. CONCLUSION AND FUTURE WORK

This paper proposes an approach to deal with conflicting norms by adding personality traits characteristics to the BDI architecture to improve the decision-making process that will decide which norms the agent shall fulfill. The main contributions of this research are: (i) include personality traits in the BDI architecture to improve the solving process of normative conflicts; (ii) implement different agent behaviors according to different personality traits, and (iii) make it possible to build software agents behaviors more similar to human behavior. As proof of concept, the approach presented in this paper can be verified by using the user scenario showed in Section V, where the agent needs to choose between bus or bicycle to go home once the weather conditions change. The emotional

BDI-agent was able to reason about the norms it would like to fulfill, and to select the plans that met the agent's intention of fulfilling, or violating, the norms.

As future work, we are deciding on an experimental study in order to complete the evaluation of our approach. Furthermore, our aim is to study other BDI architectures and platforms to investigate the possibility of extending them to support the development of emotional BDI-agents to deal with norms and normative conflicts. Last but not least, we will extend our architecture to make it possible for the BDI-agent not only to use personality traits for the solving process of normative conflicts, but also for choosing the best plans that it can execute in order to deal with the norms addressed to it.

REFERENCES

- [1] Viana, M. L. ; Paulo Alencar ; Lucena, C. . A Modeling Language for Adaptive Normative Agents. In: EUMAS, 2016, Valencia. European Conference on Multi-Agent Systems, 2016.
- [2] Figueiredo, K. Da S., Silva, V. T. Da, Braga, C. de O., Modeling norms in multi-agent systems with NormML. In Proceedings of the 6th international conference on Coordination, organizations, institutions, and norms in agent systems (COIN@AAMAS'10), Springer-Verlag, Berlin, Heidelberg, 39-57, 2010.
- [3] López, L. F.; Márquez, A. A. An Architecture for Autonomous Normative Agents, IEEE, Puebla, México, 2004.
- [4] Vasconcelos, W.W., Kollingbaum, M.J., Norman, T.J., Resolving conflict and inconsistency in norm-regulated virtual organizations. In Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems (AAMAS '07). ACM, New York, NY, USA, , Article 91 , 8 pages. 2007.
- [5] Barbosa, S. D. J., Silva, F. A. G. Da, Furtado, A. L., Casanova M. A., Plot Generation with Character-Based Decisions. Comput. Entertain. 12, 3, Article 2 (February 2015).
- [6] L. Padgham and G. Taylor, "A system for modelling agents having emotion and personality." in PRICAI Workshop on Intelligent Agent Systems, ser. Lecture Notes in Computer Science, L. Cavedon, A. S. Rao, and W. Wobcke, Eds., vol. 1209. Springer, 1996, pp. 59–71.
- [7] Bordini, R. H.; Hübner, J. F.; Wooldridge, M. Programming Multi- Agent Systemms in AgentSpeak using Jason. [S.l.]: [s.n.], 2007.
- [8] Lopez, Fabiola López. Social Power and Norms. Diss. University of Southampton, 2003.
- [9] Viana, M. L. ; Paulo Alencar ; Everton Guimarães ; Cunha, F. J. P. ; Donald Cowan ; Lucena, C. . JSAN: A Framework to Implement Normative Agents. In: SEKE, 2015, Pittsburgh. The 27th International Conference on Software Engineering & Knowledge Engineering, 2015. p. 660-665.
- [10] Bratman, Michael. Intention, plans, and practical reason. 1987.
- [11] Wooldridge, Michael; Jennings, Nicholas R.; Kinny , David. A methodology for agent-oriented analysis and design. In: In Proceedings of the Third International Conference on Autonomous Agents Agents 99. 1999.
- [12] Wooldridge, M. Ciancarini, P., Agent-Oriented Software Engineering: the State of the Art, LNCS 1957, Germany: Springer, p.45-58. 2001.
- [13] Rao, Anand S., And Michael P. Georgeff. "BDI agents: From theory to practice." ICMAS. Vol. 95. 1995.
- [14] D. Pereira, E. Oliveira, N. Moreira and L. Sarmento, "Towards an Architecture for Emotional BDI Agents," 2005 portuguese conference on artificial intelligence, Covilha, 2005, pp. 40-46.
- [15] Criado, N., Argente, E., Noriega, P., and Botti, V. Towards a Normative BDI Architecture for Norm Compliance. COIN@ MALLOW2010, pages 65–81, 2010.
- [16] dos Santos Neto, Balduino F., Viviane Torres da Silva, and Carlos JP de Lucena. "Developing Goal-Oriented Normative Agents: The NBDI Architecture." International Conference on Agents and Artificial Intelligence. Springer Berlin Heidelberg, 2011.