PUC

ISSN 0103-9741

Monografias em Ciência da Computação

n° MCC08/2017

An Architecture for Autonomous Normative BDI Agents Based on Personality Traits to Solve Normative Conflicts

Paulo Henrique Cardoso Alves

Marx Leles Viana

Carlos José Pereira de Lucena

Departamento de Informática

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO

RUA MARQUÊS DE SÃO VICENTE, 225 - CEP 22453-900

RIO DE JANEIRO - BRASIL

Monografias em Ciência da Computação, No. MCC08/2017 ISSN: 0103-9741
Editor: Prof. Carlos José Pereira de Lucena May, 2017

An Architecture for Autonomous Normative BDI Agents Based on Personality Traits to Solve Normative Conflicts *

Paulo Henrique Cardoso Alves, Marx Leles Viana,

Carlos José Pereira de Lucena

{palves, mleles, lucena}@inf.puc-rio.br

Abstract. Social norms have become one of the most promising approaches that use an informal social control to ensure a desirable social order in open multiagent systems. Normative multiagent systems offer the ability to integrate social and individual factors to provide increased levels of fidelity with respect to modeling social phenomenon such as cooperation, coordination, decision making process, organization and so on in human and artificial agent systems. 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 in the same period of time, the agent is not able to fulfill (or violate) both norms at the same time. The agent's decision about which norms to fulfill can be defined based on rewards, punishments and agent goals. Sometimes, the analysis between these attributes will not be enough to allow the agent to make the best decision. In this context, this paper introduces an architecture that considers the agent's personality traits in order to improve the normative conflict solving process. Our approach's applicability and validation is demonstrated by an experiment that reinforces the importance of considering the norms both in the agent' and society's points of view.

Keywords: Solving Normative Conflicts, Normative Agents, Multi-Agent Systems, Personality Traits

Resumo. Normas se tornaram uma das abordagens mais promissoras que utilizam o controle social para auxiliar o convívio em sociedade em um sistema multiagentes. Sistemas multiagentes normativos oferecem a habilidade para integrar fatores sociais e individuais para promover o aumento nos níveis de fidelidade a respeito de fenômenos sociais como cooperação, coordenação e tomada de decisão. Entretanto, eventualmente algumas normas podem entrar em conflito - por exemplo, quando existe uma norma que proíbe um agente de realizar uma ação em particular e outra norma que obriga o mesmo agente a realizar a mesma ação no mesmo intervalo de tempo. A decisão do agente sobre quais normas serão cumpridas pode ser definida com base nas recompensas, punições e objetivos do agente. Porém, a avaliação desses atributos poderá não ser o suficiente para permitir que o agente efetue a melhor tomada de decisão. Nesse contexto, essa monografia analisa os resultados do experimento desenvolvido com diferentes abordagens para lidar com conflitos em sistemas normativos e para avaliar como cada perfil se comporta em determinadas situações.

Palavras-chave: Resolução de Conflitos Normativos, Agentes Normativos, Sistemas Multiagentes, Traços de Personalidade.

* This work has been sponsored by the Ministério de Ciência e Tecnologia da Presidência da República Federativa do Brasil (CAPES)

In charge of publications:

Rosane Teles Lins Castilho
Assessoria de Biblioteca, Documentação e Informação
PUC-Rio Departamento de Informática
Rua Marquês de São Vicente, 225 - Gávea
22453-900 Rio de Janeiro RJ Brasil
Tel. +55 21 3114-1516 Fax: +55 21 3114-1530
E-mail: bib-di@inf.puc-rio.br

Table of Contents

1 Introduction

Multiagent Systems (MASs) are societies in which these heterogeneous and individually designed entities (agents) work to accomplish common or independent goals (VIANA,ALENCAR AND LUCENA, 2016). 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 to which the agents are obligated (agents must accomplish a specific outcome), permitted (agents can act in a particular way) or prohibited (agents must not act in a specific way), to encourage the fulfillment of the norm through rewards definition and discouragement of norm violation by pointing out the punishments (FIGUEIREDO, SILVA and BRAGA, 2010).

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 (VIANA,ALENCAR AND LUCENA, 2016). 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 (LÓPEZ and MÁRQUEZ, 2004) 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 (MCCRAE and JOHN, 1992). 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 to permit the agent to make the best decision.

The abstract normative agent architecture developed by (LÓPEZ and MÁRQUEZ, 2004), has four main steps: (i) agent perception, i.e., when the agent's beliefs and a set of norms are updated; (ii) norm adoption, i.e., when agents verify which norms are addressed to them; (iii) norm deliberation, i.e., when agents verify which norms they intend to fulfill, or reject, and (iv) norm compliance, i.e., when agents verify which norms they will comply with. 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 make some different decisions involving personality traits based on OCEAN model (MCCRAE and JOHN, 1992), setting a weight for each one of these characteristics. We will present an experiment comparing different approaches to deal with normative conflicts based on social profiles and personality traits. This will illustrate the new deliberation process proposed in this paper.

In this context, we present an approach to build BDI agents with personality traits (Belief-Desire-Intention) that also considers other agent personality traits (BARBOSA, SILVA, FURTADO and CASANOVA, 2015) to improve the decision-making process for solving normative conflicts. This approach aims at offering 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. By using these new functions, it is possible to build agents that: (i) use personality traits to improve the solution among normative conflicts, (ii) implement the agent's behavior similar to a human's behavior, and (iii) evaluate the effects on its desires with respect to the fulfillment or violation of a norm and use all of these points to conduct experiments to learn how different strategies could change an agent's behavior.

The paper is structured as follows. Section 2 focuses on the background, while Section 3 discusses related work. Section 4 presents the personality traits BDI approach to solve normative conflicts. Section 5 presents some simulations that help to evaluate our approach. Finally, Section 6 shows our conclusions and future work.

2 Background

This section describes the main concepts related to agents and multiagent systems. First, we will discuss norms and BDI (Belief-Desire-Intention) architecture. We will also discuss the relation between normative conflicts.

2.1 Norms

Norms are designed to regulate the agent's behavior, and therefore, a norm definition should include the address of the agent being regulated (BORDINI, HÜBNER and WOOLDRIDGE, 2007). 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 (VIANA et al., 2015), which is composed by the representation of the element *norm* – it contains many different properties. These properties are briefly described in Table I. For example, the property *Addressee* is used to specify the agents or roles responsible for fulfilling the norm.

Property	**Description**
Addressee	It is the agent or role responsible for fulfilling the norm.
Activation	It is the activation condition for the norm to become active.
Expiration	It is the expiration condition for the norm to become inactive.
Rewards	It represents the set of rewards to be given to the agent to fulfill a norm.
Punishments	It is the set of the punishments to be given to the agent for violating a norm.
DeonticConcept	It indicates if the norm states an obligation, a permission or a prohibition.
State	It describes the set of states being regulated.

Table 1: Norm Description

In order to better understand the application of norms to regulate agents with a different social profile, we made a comparison for showing how many normative goals and agent goals are complied and fulfilled for each single approach. Furthermore, to better understand the definition of norms and their representation, imagine a user scenario where the employee agent has to decide the transportation type to go home. The agent's goal is to increase physical conditioning and has the following options to go home: (i) by bicycle, which is a way to satisfy the agent's goal, and (ii) by bus, if it is raining; thus, the agent cannot accomplish its goal at this time.

In addition, each employee agent should decide according to specific norms. Eventually, a norm is sent to each employee agent with the following state: "go home by bus, it is raining". This norm has the following attributes: (i) addressees are employee agents; (ii) the required deontic concept is prohibition, because it prohibits the agent to go to home by bicycle, and (iii) when an agent agrees to a norm, it will receive a reward. In this case, the reward may be not get the flu. If the employee agent violates the norm, the agent will receive the punishment associated with the norm. For example, when it is raining and the employee agent does not want to work out so badly, it then will violate the norm by going home by bicycle, which will result in the decrease of the agent's health, because the agent will probably come down with the flu. In this case, a punishment associated with the norm will be applied to the agent, i.e., the agent cannot work the next day because it is sick. Note that the norm is activated when it is raining. In turn, the norm expires when the weather is sunny.

2.2 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 (VASCONCELOS, KOLLINGBAUM and NORMAN, 2007). 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. The agent can realize any action in the environment until an active norm restricts its goals. For example, Figure 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.

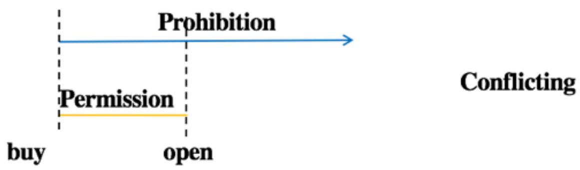

Figure 1. Conflict - Prohibition and Permission.

Figure 2 presents another scenario of conflicting norms - the seller agent can only reprice the products before the store opens and another norm permits the seller agent to reprice them when the store is open and there is a sale.

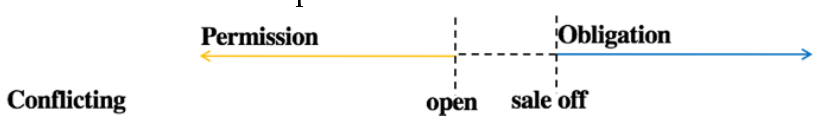

Figure 2. Conflict - Permission and Obligation.

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

2.3 BDI Architecture

The BDI (*Belief Desire Intention*) model was proposed by (BRATMAN, 1987) as a philosophical theory of practical reasoning, representing, respectively, the agent's information, motivational and deliberative states. There are two main steps: (i) apply a filter to make a set of goals to which the agent must 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 (WOOLDRIDGE and CIANCARINI, 2001).

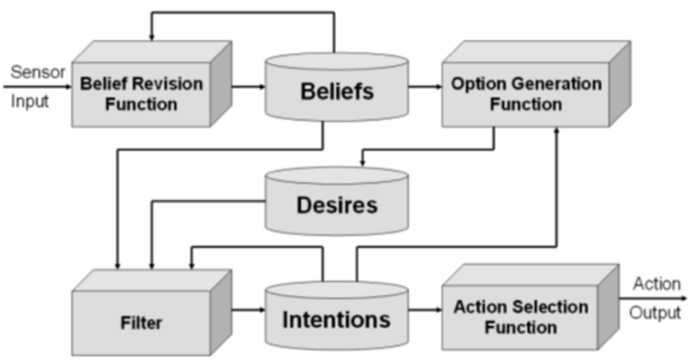

Figure 3. Generic BDI architecture (WOOLDRIDGE, JENNINGS and KINNY, 1999).

The BDI model is composed of 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. Figure 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 and Georgeff (RAO and GEORGEFF, 1995) 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 all together, 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.

2.4 Personality Traits

The big-5 model (MCCRAE and JOHN, 1992)., also knew as OCEAN model, provides a mechanism to define personality traits based on such concepts and defines five factors: (i) Openness, describing a dimension of personality that portrays the imaginative, creative aspect of human character, (ii) Conscientiousness, determining how much an individual is organized and careful, (iii) Extroversion, related to how outgoing and sociable a person is, (iv) Agreeableness, which is about friendliness, generosity and the tendency to get along with other people, and (v) Neuroticism, referring to emotional instability and the tendency to experience negative emotions.

Each factor is composed of many traits, which basically are used to describe people (MCCRAE and JOHN, 1992) (GOLDBERG, 1990). The presented factors will be used to help the agent decision-making process and also in plan selection according to agent individual goals and norms intended.

Based on OCEAN model, the personality traits may be built through the distribution of weights between the factors: (i) Openness to experience; (ii) Conscientiousness; (iii) Extraversion; (iv) Agreeableness, and (v) Neuroticism. In the Figure 4, agent 1 may be creative and adventurous, however, agent 2 may be careful.

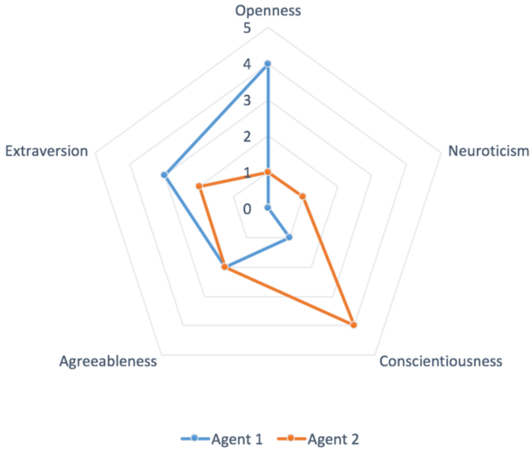

Figure 4. OCEAN model application example

3 Related Work

This section describes some related work: (i) the solution for normative conflicts (LOPEZ, 2003), (CRIADO, ARGENTE, NORIEGA and BOTTI, 2010), (NETO, SILVA and LUCENA, 2011); (ii) architecture designs considering the agent's emotional state (PEREIRA, OLIVEIRA, MOREIRA and SARMENTO, 2005), and (iii) the agent's personality (BARBOSA, SILVA, FURTADO and CASANOVA, 2015), (PADGHAM and TAYLOR, 1997) , (JONES, SAUNIER and LOURDEAUX, 2009).

Pereira et al. (PEREIRA, OLIVEIRA, MOREIRA and SARMENTO, 2005) proposed an architecture based on the BDI (Belief-Desire-Intention) model to support artificial emotions, including internal representations of the agent's capabilities and resources. This research introduces subjects, such as artificial emotions, agent means and BDI architecture. Furthermore, a common-sense definition of new mental states, such as emotions, was developed and made them influence the BDI architecture through the common-sense understanding of the way they positively affect human reasoning. The authors defined a new concept: Fear, an informational data structure that reports situations which an agent should avoid. This work presents the Emotional-BDI architecture as an extended version of the classic BDI. However, the authors do not compare the results with other approaches, which apply or do not apply emotions and (PEREIRA, OLIVEIRA, MOREIRA and SARMENTO, 2005) do not provide support to solve normative conflicts.

The authors in (BARBOSA, SILVA, FURTADO and CASANOVA, 2015) built a decision process to work as part of the story-telling systems wherein narrative plots emerge from the acting characters' behaviors 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.

Padgham and Taylor (PADGHAM and TAYLOR, 1997) present an architecture considering an analysis of human behavior based on emotions. The authors present how emotions and personality traits interact with goal-oriented behavior and describe some simplifications that were made in order to build an initial interactive environment for experimentation with animated agents simulating personality and emotions. According to the authors, emotions can affect behavior, directly or indirectly. An effect includes such things as re-prioritizing, adding and deleting goals. For example, an agent experiencing gratitude, might delete, or give low priority to those goals that conflict with its grateful goals. In addition, this work presents personality as a notion of factors that may affect agents. For instance, a happy agent moves faster, and with more bounce, while a sad agent is slower and with less spring to its movements. The authors also developed two scenarios. In both, agents successfully displayed the features of different personalities based on their emotional profiles. However, this work does not deal with norms but, rather, just applies the emotions context solely to the agent's goals. Part of our research was based on this work, which allowed us to add more human attributes, thus improving our conflicting norms resolution.

Jones et al. (JONES, SAUNIER and LOURDEAUX, 2009) developed a BDI extension to consider physiology, emotions, and personality. It is used to model crisis situations; for instance, terrorist attacks. The emotions were used in pairs such as fear/ hope, anger/gratitude and shame/pride. The physiology may be affected by the simulation environment and may change the agent's health. The following characteristics were considered: stress, hunger/thirst, temperature fatigue, injuries and contamination. The personality is a set of characteristics that determines that agents are psychologically, mentally and ethically different from each other. However, this approach was not applied in Normative Multiagent Systems to evaluate different behaviors that may emerge with personality traits application.

Some approaches (LOPEZ, 2003), (CRIADO, ARGENTE, NORIEGA and BOTTI, 2010), (NETO, SILVA and LUCENA, 2011) 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. (CRIADO, ARGENTE, NORIEGA and BOTTI, 2010) 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 neither supports a strategy nor considers the agent's personality traits to deal with conflicts between norms.

Lopes et al. (LOPEZ, 2003) defined a set of strategies that can be adopted by agents to deal with norms as follows: Pressured, Rebellious and Social. For example, the Pressured strategy occurs when agents fulfill the norms to achieve their individual goals considering only the punishments that will harm them. Another is the Rebellious strategy, in which agents consider only their individual goals and violate all the environment's norms. Finally, the Social strategy happens when agents first of all comply with norms and after verify if is possible to fulfill some individual goals. 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 multiagent systems by including new strategies. In addition, this work can neither extend mechanisms to collect information during the simulations nor extend mechanisms to generate norms and agent goals. Furthermore, the agent cannot detect and overcome normative conflicts.

Finally, Santos Neto et al. (NETO, SILVA and LUCENA, 2011) propose the NBDI architecture, based on the Criado et al. (CRIADO, ARGENTE, NORIEGA and BOTTI, 2010) research, to develop goal-oriented normative agents whose priority is the accomplishment of their own desires while evaluating the pros and cons associated with the fulfillment or violation of the norms. To make this possible, the BDI architecture was extended by including norms related functions to check incoming perceptions, select norms based on the agent's desires and intentions. A detection conflict and a solving conflict algorithm were developed based on norms contributions; in the case of conflicts between norms, the one with the highest contribution to the achievement of the agent's desires and intentions can be selected. If the norm contributions have equal values, then the first norm will be selected. Therefore, as it is possible to observe, sometimes the norm contribution is not enough for the agent to make a better decision. We identified this gap and improved the decision-making process, adding the personality traits concepts.

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 and improve the deliberation process. 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.

4 BDI Agents with Personality Traits: An approach

This section describes the main concepts required to understand the approach based on BDI agents with personality traits used to improve the solution of normative conflicts and, after helping the deliberation process, to deal with non-conflicting norms and the agent goals. In addition, we provide a software framework overview and discuss its different components.

4.1 The Architecture

The Personality Traits 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 conflicts resolution and normative deliberation process. The architecture foundation was based on the abstract normative agent architecture developed in (LÓPEZ and MÁRQUEZ, 2004). Figure 5 presents our BDI agent with personality traits architecture to solve normative conflicts.

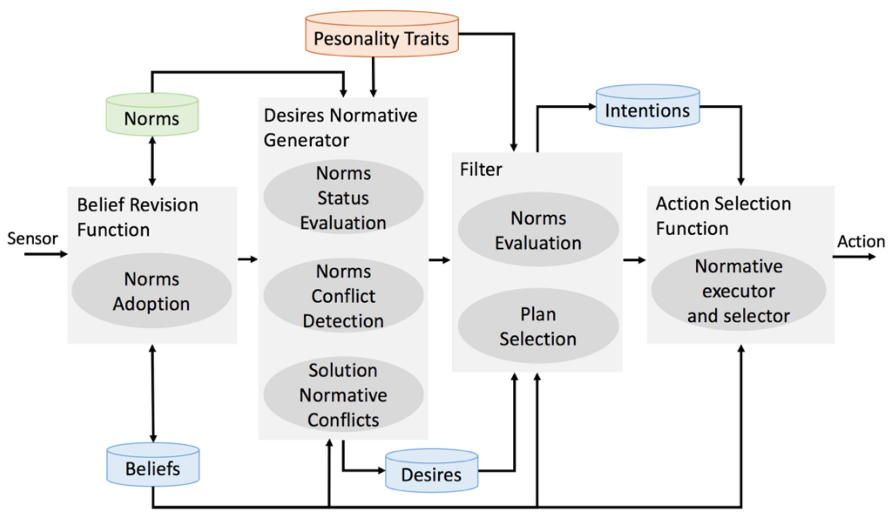

Figure 5. Internal architecture of the BDI Agent based on Personality Traits to Improve Normative Conflicts Solution.

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 agent's 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 of 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 based on OCEAN model. Table 2 shows the personality traits that we consider – drives, attitudes and emotions as in (MCCRAE and JOHN, 1992). As a result, a set of non conflicting norms are exported to the next step. These norms are the agent's Desires.

Drives	**Attitudes**	**Emotions**
Sense of duty	Careful	Anger
Material gain	Adaptable	Fear
Spiritual endeavor	Self-controlled	Surprise

Table 2: Personality Traits Example

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*, composed of 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 improve the normative conflict solving process, considering personality traits inserted into the BDI reasoning process.

4.2 The Framework

Inspired by the JSAN architecture (VIANA, ALENCAR and LUCENA, 2016), which uses different norm strategies to deal with the norm, taking into account the different agent's social levels, as in (LOPEZ, 2003), we built a new approach by introducing personality traits aiming to improve the solution of the normative conflict. Our framework provides the decision-making process described in Section 4.1. Figure 6 shows the framework architecture..

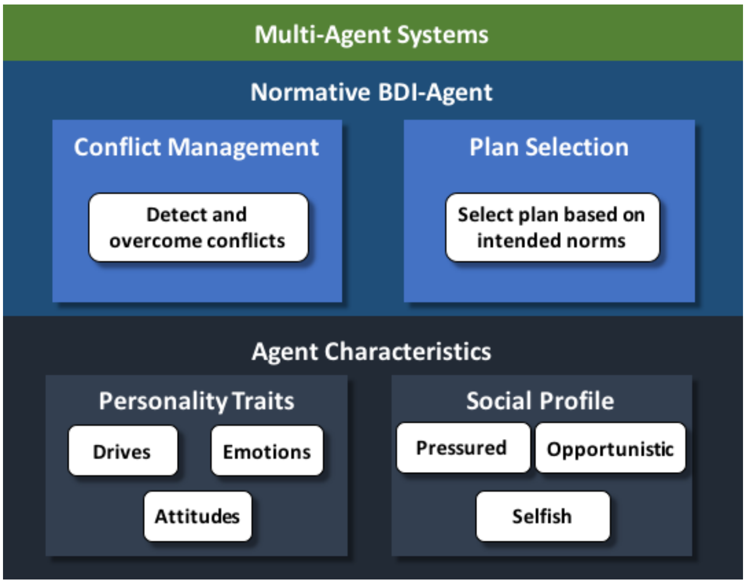

Figure 6. The Framework Architecture.

The NormativeBDI-Agent class is composed of goals, role, norms, beliefs, desires, intentions, and personality traits. By using these attributes, the agent starts the decision making process to solve normative conflicts. By means of the normative conflict solving process, the agent will choose the norms that it will add to the Intentions set and finally will decide which norms will be fulfilled according to the agent's social profile, as in (BORDINI and WOOLDRIDGE, 2007) and (LOPEZ, 2003).

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 of 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 analyzes 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 their 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.

4.3 Experiment

Our initial experimentation includes four kinds of agents to deal with norms, such as (LÓPEZ and MÁRQUEZ, 2004) and (NETO, SILVA and LUCENA, 2011). The (LÓPEZ and MÁRQUEZ, 2004) approaches deal with norms considering the strategies: (i) Social, i.e., the agent complies with all of active norms directed to him and after it verifies which goals can be fulfilled, if there are conflicts, randomly selects one norm of each conflicting norms set to be complied with, (ii) Rebellious, i.e., the agent violates all norms and fulfills all goals, and in this case it does not matter if there are conflicting norms; the agent never will comply with any norms, and (iii) Pressured, i.e., the agent fulfills its goals first and then verifies which norms can be complied with; if there are conflicting norms, the agent will measure the goals' importance to decide which norm it will fulfill. The approach developed by (NETO, SILVA and LUCENA, 2011) considers the normative contribution generated by evaluation between: (i) the norms' rewards and punishments, and (ii) the goals' importance.

We choose these examples to compare with our approach because they represent the most common strategies followed by agents when they face a norm compliance decision. The approach presented in this paper was based on (NETO, SILVA and LUCENA, 2011) and improved by personality traits characteristics. We consider the normative contribution adding the weight of the personality traits. When a conflict is identified, i.e., there are at least two norms with: (i) opposite deontic concept, and (ii) both norms are active, the norm contribution is evaluated for each one of conflicting norms and there are a few steps to follow: (i) for each goal, the goal importance is increased by a personality traits weight, (ii) for each goal allowed by a norm, i.e., the norm does not restrict this goal, the norm contribution is increased, adding the goal importance, and (iii) for each norm that is active at the same time and has opposite deontic concept, the norm with the better norm contribution value is selected.

For the non-conflicting norms (i) a set of norms indexed by the goals that the norm restrict is created, (ii) for each non-conflicting norm the norm contribution is increased adding the norm contribution value for each norm in this set that restrict the same goal, (iii) the norms contribution and goals increased by personality traits are evaluated, and (iv) the better value is selected and this norm or goal is selected to be fulfilled. Our interest here is to observe how both the social contribution, and the individual satisfaction of each agent change according to both the strategy for norm compliance it chooses and the increase in the number of conflicts between the norms it has to play with and its personal goals. The social contribution of an agent is defined as the number of times the agent has complied with its responsibilities (i.e., the number of times the agent has fulfilled the norms) related to the number of norms that become active. The individual satisfaction of an agent is the number of goals achieved related to the number of goals generated.

The experiment using all of these different approaches was developed and some variables are fixed as follow. First, a base of goals to represent all the goals that an agent might have is randomly created. Second, a motivation value is associated to each goal in this set to represent the importance of each goal. In addition, each goal might have a personality trait associated, so if there is an agent that has this personality trait, this goal will be increased by the personality trait value. Both punishments and rewards in each norm are also randomly generated, as also are the deontic concept and activation time. Thus, norms are evaluated by agents following different strategies so that similar inputs produce different outcomes.

In our experimentation, we recorded both the social contribution and the individual satisfaction of each agent for a particular percentage of conflicts over a period of time. First, no conflicts are considered, meaning that all norms and goals could be fulfilled. Then the experiment was repeated, with the number of conflicts increased in a proportion of 25% until all norms conflict themselves. Each experiment consisted of 100 runs, and in each run, 10 goals and norms were used.

First, the Pressured approach shows that the agent achieves more individual goals than contributes to the society. Figure 7 shows the agent's behavior in different conflicting norms situations.

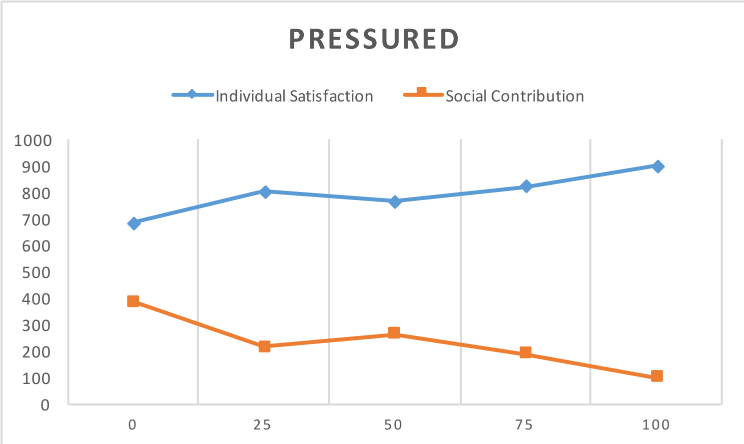

Figure 7. Pressured strategy.

The Social approach shows that initially, with no conflicting norms, the agent complies with all norms because, first, the agent complies with all active adopted norms and then decides which goals will be selected to fulfill. Moreover, the number of goals achieved increases gradually, as shown in Figure 8.

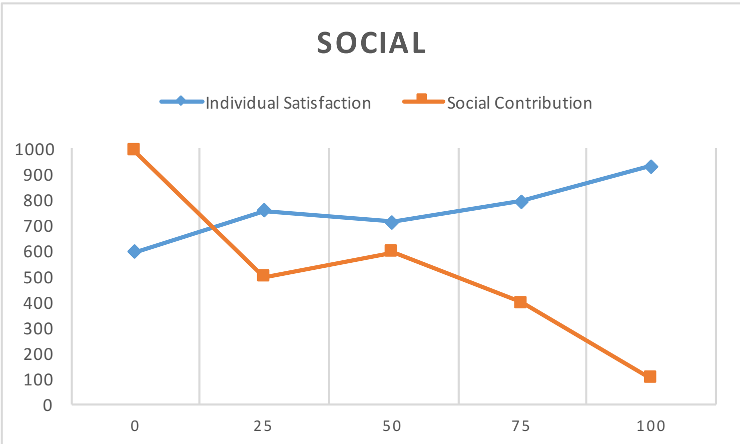

Figure 8. Social strategy.

Agents with the Rebellious strategy violate all norms and, as no goal will be restricted, all of them will be achieved. Figure 9 shows this behavior; it is important to notice that the rewards and the punishment values are not taken into account. In this situation, the agent always will receive a punishment for violate norms that restrict its goals.

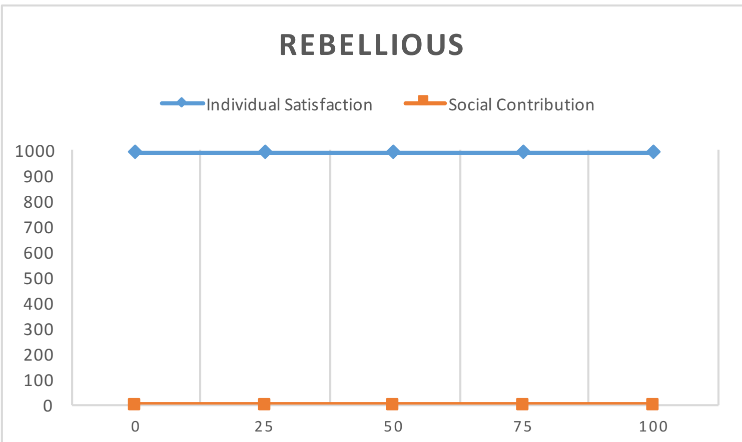

Figure 9. Rebellious strategy.

The agent with NBDI strategy considers the evaluation of contribution for fulfilling or violating each norm before deciding for fulfilling or violating it. Figure 10 shows the goals and norms pattern; the more conflicting the norms, the more the goals are achieved.

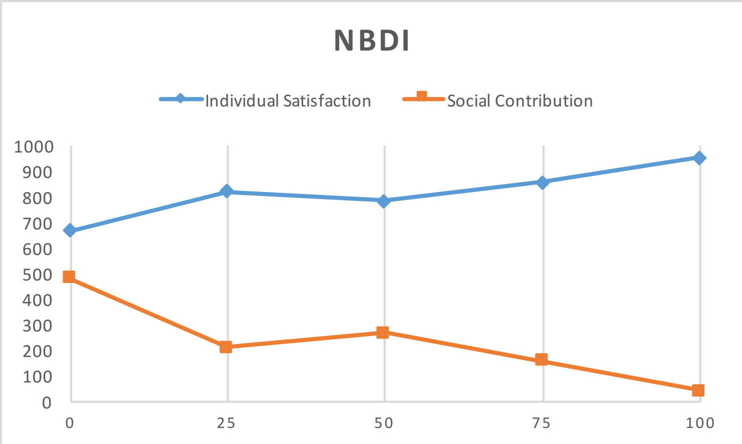

Figure 10. NBDI strategy.

Personality Traits strategy considers the norm contribution developed in NBDI strategy and increase this value with personality traits. The experiment results are similar to the NBDI strategy, although the agent meets more individual goals. Figure 11 shows the agent behavior regarding norms complience and goals achievement.

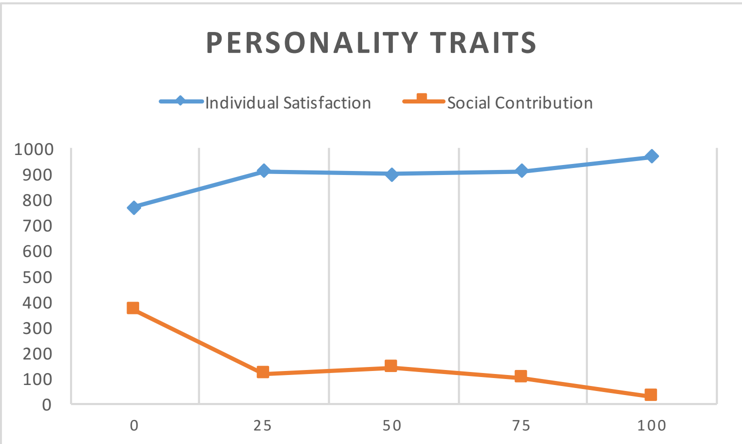

Figure 11. Personality Traits strategy.

As can be observed, the personality traits approach encourages the agent to fulfill its goals and, if there is a personality trait with a null value, the performance will be the same as presented by NBDI. The greater the weight of the personality traits, the higher will be the number of individual goals.

Figure 12 shows the comparison between all of five strategies and the Personality Traits strategy to achieve more goals than Social, Pressured and NBDI strategies. It shows that the Personality Traits approach helps the agent to fulfill more individual goals and increases the individual satisfaction.

Figure 12. Individual Satisfaction Resume.

Figure 13 shows all of the five different strategies, comparing the social contribution between them. As can be observed, strategies that achieve more goals comply with fewer norms; therefore, the Personality Traits strategy complies with fewer norms than other strategies, except the Rebellious strategies, which violate all norms always.

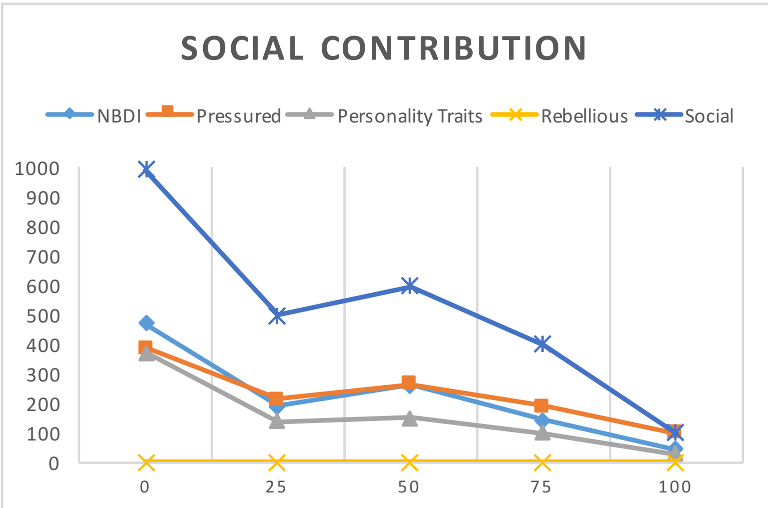

Figure 13. Social Contribution Resume.

5 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 agent behaviors that are more similar to human behavior to improve experiments where this factor is relevant. The personality traits 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. Moreover, the experiment developed showed that the Personality Traits strategy results were similar to the NBDI strategy, although the agent with personality traits chooses to achieve more goals than with the other strategies.

As future work, we are deciding on an experimental study in order to apply fuzzy logic to deal with changes found in the real world, such as the chance to become sick, if you stay in the rain. Furthermore, the punishment for becoming ill is also variable. An agent's punishment may range from sneezing to pneumonia. The severity of the illness could be a factor of the agent's current health state and how fast to recovery may also be part of the agent's personality profile. So, when the agent must decide whether to ride the bike in the rain, it must calculate the reward (fitness gained) against the possibility of becoming sick (may or may not get sick) and the consequences (punishment) that could range from very mild (sneezing) to very serious (pneumonia). We also plan to implement this approach in other more complex scenarios, taking into account personality traits. For example: (i) people in risk areas, where firefighters are responsible for planning their evacuation, and (ii) crime prevention, where the police are responsible for arresting criminals and keeping civilians safe. Last but not least, we will apply these different strategies in environments with more agents, to analyze their behavior, evaluating norms and internal goals in a society.

References

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).

BORDINI, R. H.; HÜBNER, J. F.; WOOLDRIDGE, M. **Programming Multi- Agent Systems in AgentSpeak using Jason**. [S.l.]: [s.n.], 2007.

BRATMAN, M.. **Intention, plans, and practical reason**. 1987.

CRIADO, N.; ARGENTE, E.; NORIEGA, P; BOTTI, V., **Towards a Normative BDI Architecture for Norm Compliance**. COIN@ MALLOW2010, pages 65–81, 2010.

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.

GOLDBERG, L.R., **An Alternative "Description of Personality": The Big-Five Factor Structure**. Journal of Personality and Social Psychology, 59 (6). 1216-1229, 1990.

JONES, H., SAUNIER, J., LOURDEAUX, D., **Personality, Emotions and Physiology in a BDI Agent Architecture: The PEP -> BDI Model**, IEEE, 263-266, 2009.

LOPEZ, F. L.. **Social Power and Norms**. Diss. University of Southampton, 2003.

LÓPEZ, L. F.; MÁRQUEZ, A. A. **An Architeture for Autonomous Normative Agent**s, IEEE, Puebla, México, 2004.

MCCRAE, R., John, O., **An introduction to the five-factor model and its applications, Journal of Personality**, vol. 60, pp. 175–215, 1992.

NETO, B. dos S.; SILVA, V. T. da; LUCENA, C. J. P de, **Developing Goal-Oriented Normative Agents: The NBDI Architecture,** International Conference on Agents and Artificial Intelligence. Springer Berlin Heidelberg, 2011

PADGHAM, L.; TAYLOR, G., **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.

PADGHAM, L.; TAYLOR, G., **A system for modelling agents having emotion and personality**. In: Intelligent Agent Systems Theoretical and Practical Issues. Springer Berlin Heidelberg, 1997. p. 59-71.

PEREIRA, D.; OLIVEIRA, E.; MOREIRA, N.; SARMENTO, L., **Towards an Architecture for Emotional BDI Agents**, 2005 Portuguese conference on artificial intelligence, Covilha, 2005, pp. 40-46.

RAO, A. S.; GEORGEFF, M. P., **BDI agents: From theory to practice,** ICMAS. Vol. 95. 1995.

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.

VIANA, M. L.; ALENCAR, P.; GUIMARÃES, E.; CUNHA, F. J. P. ; COWAN, D.; 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.

VIANA, M. L.; ALENCAR, Paulo; LUCENA, C. **A Modeling Language for Adaptive Normative Agents**. In: EUMAS, 2016, Valencia. European Conference on Multi-Agent Systems, 2016.

Wooldridge, M. Ciancarini, P., Agent-Oriented Software Engineering: the State of the Art, LNCS 1957, Germany: Springer, p.45-58. 2001

WOOLDRIDGE, M.; JENNINGS, N. R.**, Pitfalls of agent-oriented development,** Proceedings of the Second International Conference on Autonomous Agents (Agents'98), ACM Press, pp. 385-391, 1998.

WOOLDRIDGE, M.; JENNINGS, N. R.; KINNY, D., **A methodology for agent-oriented analysis and design**. In: Proceedings of the Third International Conference on Autonomous Agents 99. 1999.