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

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

Keywords: Solving Normative Conflicts, Normative Agents, Multiagent Systems, Personality Traits.

Abstract: Norms are promising mechanisms of 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 modelling social phenomena such as cooperation; coordination; decision-making process, and organization in 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, the agent is not able to fulfill both norms at the same time. The agent's decision about which norms to fulfill can be defined based on rewards, punishments and agent's goals. Sometimes, the analysis between these attributes will not be enough to allow the agent to make the best decision. This paper introduces an architecture that considers the agent's personality traits in order to improve the normative conflict solving process. In addition, the agent can execute different behaviors with equal environment variables, just by changing its own internal characteristics. The applicability and validation of our approach are demonstrated by an experiment that reinforces the importance of the society's norms.

1 INTRODUCTION

Multiagent Systems (MASs) are societies in which these heterogeneous and individually designed entities (agents) work to accomplish common or independent goals (Viana et al., 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: (i) obligated (agents must accomplish a specific outcome); (ii) permitted (agents can act in a particular way) or (iii) prohibited (agents must not act in a specific way) all to encourage the fulfillment of the norm through rewards definition and discouragement of norm violation by pointing out the punishments (Figueiredo et al, 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 environment 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 et al., 2016). Since an agent's priority is the satisfaction of its own goals, before complying with the 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 punishments are the means by which the agents 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 punishments 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

Alves, P., Viana, M. and Lucena, C.
An Architecture for Autonomous Normative BDI Agents based on Personality Traits to Solve Normative Conflicts.
DOI: 10.5220/0006599300800090
In *Proceedings of the 10th International Conference on Agents and Artificial Intelligence (ICAART 2018) - Volume 1*, pages 80-90
ISBN: 978-989-758-275-2

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 violate, 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 the 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.

Within this context, we present an approach that builds BDI agents with personality traits (Barbosa et al., 2015) to improve the decision-making process for the solution of 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 in order 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 between normative conflicts, and (ii) evaluate the effects on its desires with respect to the fulfilment, or violation, of a norm and thus use all of these functions 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 BDI-agent approach to personality traits to solve normative conflicts. Section 5 presents the experiment that evaluates our approach. Finally, Section 6 shows our conclusion 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 behavior of the agent, and therefore, a norm definition should include the address of the agent being regulated (Bordini et al., 2007). However, norms are different from laws, and they cannot force agents to comply with them. Agents are autonomous entities, so norms can only suggest and present the expected behavior to which the agent will decide to comply with, or not.

In this work, we used the norm representation described in (Viana et al., 2015) . Norms properties are briefly described in Table 1. For example, the property Addressee is used to specify the agents or roles responsible for fulfilling the norm.

Table 1: Norm Description.

Property	Description
Addressee	It is the agent or role responsible for fulfilling the norm
Activation	It is the condition for the norm to become active
Expiration	It is the 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 punishments to be given to the agent for violating a norm
Deontic Concept	It indicates if the norm states an obligation, a permission or a prohibition
State	It describes the set of states being regulated

In order to better understand the application of norms to regulate agents with a different social profile, we made a comparison between the social contribution and the individual satisfaction of the agent for fulfilling, or violating, the norms for each 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, in which case, 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 statement: "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 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 getting 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 really wants to work out, 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 obligated (Vasconcelos et al., 2007). These conflicts and inconsistencies may be caused by a norm that prohibits an agent to perform a particular action while another norm obligates 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 bring back the product bought and at the same time another norm defines that the buyer agent can return the product bought before opening it.

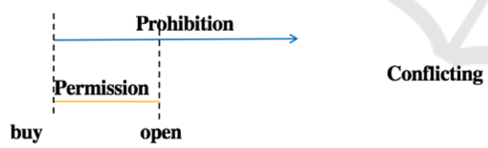

Figure 1: Conflict - Prohibition and Permission.

Figure 2 presents another scenario of conflicting norms — the seller agent can only re-price the products before the store opens and another norm permits the seller agent to re-price 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 important 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, the information, the motivational and deliberative states of the agent, respectively. There are two main steps: (i) applying a filter to make a set of goals to which the agent must commit to serve as the basis of its beliefs, and (ii) finding a way to understand how the desires produced can be fulfilled based on the agent's available resources (Wooldridge and Ciancarini, 1999).

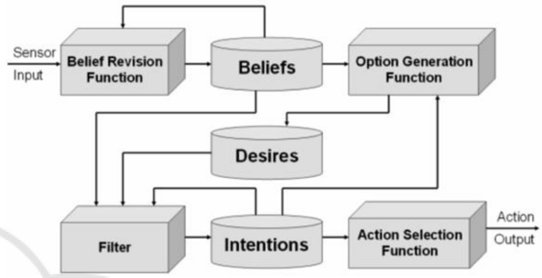

Figure 3: Generic BDI architecture (Wooldridge et al., 1999).

Figure 3 shows the BDI model, which is composed of three mental states: (i) beliefs, which represent the environment factors that are updated after each perceived action — they represent the agent's 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.

The 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 functionality is the means to the 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, 2011), also known 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 the 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, 2011) (Goldberg, 1990). The factors presented will be used to help the agent's decision-making process and plan selection, according to the agent's individual goals and intended norms.

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

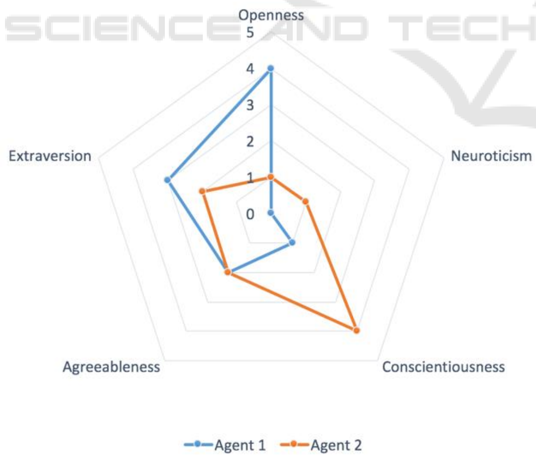

Figure 4: OCEAN model application example.

3 RELATED WORK

This section describes some related work: (i) the solution for normative conflicts (López, 2003), (Criado et al., 2010), (Neto et al., 2011); (ii) architecture designs considering the agent's emotional state (Pereira et al., 2005), and (iii) the agent's personality (Barbosa et al., 2015), (Jones et al., 2009).

Pereira et al. (Pereira et al., 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 influenced 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 Personality Traits in BDI-Agent architecture as an extended version of the classic BDI. However, the authors do not compare the results with other approaches that may, or may not, apply emotions and neither provide support to solve normative conflicts (Pereira et al., 2005).

The authors in (Barbosa et al., 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 to MASs, which creates an opportunity to improve the agent's decision-making process to deal with normative conflicts.

Jones et al. (Jones et al., 2009) developed a BDI extension to consider physiology, emotions, and personality. It is used to model crisis situations such as terrorist attacks, for instance. 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 applications.

Some approaches (López, 2003), (Criado et al., 2010), (Neto et al., 2011) have been proposed in the literature to develop agents that evaluate the effects of solving normative conflicts. For instance, the n-BDI architecture defined by Criado et al. (Criado et al., 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.

Lopez et al. (López, 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 it is possible to fulfill some individual goals. Although this work provides some mechanisms for the agents to collect norms, the authors do not 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 et al., 2011) propose the NBDI architecture, based on the Criado et al. (Criado et al., 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, and 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 concept.

As none of this related work deals with norm 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 PERSONALITY TRAITS IN BDI-AGENTS

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

4.1 The Architecture

The Personality Traits in BDI agents approach that can solve the normative conflicts were inspired on the concepts presented in the background and the related work sections.

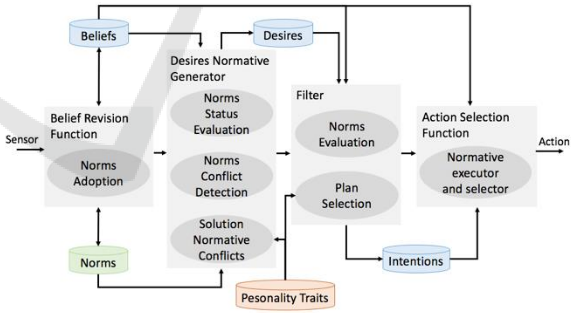

Figure 5: The architecture.

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.

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 change the agent's behavior, 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 are described below.

The first step involves the agent's perception in the Belief Revision Function, where the agent perceives the active norms in the environment addressed to it by means of its 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 the OCEAN model. Table 2 shows some examples of personality traits composition that we consider: drives, attitudes and emotions, as in (Mccrae and John, 1992). Our personality traits model has only two properties: (i) a name and (ii) a value indicating its weight.

Table 2: Personality Traits Examples.

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

The norms analyses are based on the normative contribution, which is composed by the evaluation between rewards, punishments, goals and personality traits. Figure 6 shows the normative contribution equation.

$$NC_i = NReward_i + NPunishment_i - \sum_{j=0}^{m} NPunishments_j$$

$$NC_i = NC_i + \sum_{j=0}^{p} \sum_{k=0}^{q} (AgentGoal_j + AgentPersonalityTrait_k)$$

Figure 6: Normative Contribution Equation.

The normative contribution concept was extended from (Neto, 2011). We added the goals and personality traits weights. The *m* bound refers to the summation of the activated norms addressed to the agent; the *p* bound refers to the summation of the agent's goals, and *q* refers to the summation of the agent's personality traits. Table 3 describes the goal properties.

Table 3: Goal Properties.

Property	Description
Name	It is the name of the goal
Value	It is the value that represents the importance of this goal
Norm Required	It is the set of the norms required to permit that the goal be accomplished
Personality Trait	It is the set of the personality traits required to permit that the goal be accomplished
Belief Required	It is the set of the beliefs required to permit that the goal be accomplished

The personality traits are used only in two situations: (i) at Solution Normative Conflicts through the equation shown in Figure 6 and (ii) at the plan selection step. Table 4 describes the plan properties. A set of non-conflicting norms is exported to the next step. The goals that are not restricted by the norms are the agent's Desires.

Table 4: Plan Properties.

Property	Description
Name	It is the name of the goal
Value	It is the value that represents the importance of this plan
Personality Trait	It is the set of the personality traits that contribute to this plan execution
Goal Required	It is the set of the goals required to permit that the plan be activated

The third step is the Normative Filter, which is composed of two processes taking into account the agent's personality traits: (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 will choose its best plans in the Intentions set.

Finally, the fourth step is the Action Selection function, which is composed of the Normative executor and selector. 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 to improve the normative conflict solving process, considering personality traits inserts into the BDI reasoning process.

4.2 The Framework

Inspired by the JSAN architecture (Viana et al., 2015), which uses different normative strategies to deal with norms and takes into account the different agent's social levels, as in (López, 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 7 shows the framework architecture.

The Normative BDI 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. In 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 et al., 2007) and (López, 2003).

The solving process of normative conflicts starts with the calculation of each norm's normative contribution, wherein the agent evaluates its rewards and punishments and compares each normative contribution with other 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 choice of the normative goals that can be fulfilled according to the agent's goals and its personality traits.

Figure 7: The Framework architecture.

The agent will verify which goal 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 achieved 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 one norm that obligates an 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.

5 EXPERIMENT

Our initial experimentation includes different kinds of agents to deal with norms, such as described in (López and Márquez, 2004) and (Neto et al., 2011). The (López and Márquez, 2004) approaches deal with norms considering the following strategies: (i) Social, i.e., the agent fulfills all of the active norms addressed to it and then it verifies which goals can be fulfilled; if there are conflicts, it randomly selects one norm from 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 will never fulfill any norms, and (iii) Pressured, i.e., the agent only fulfills the norms whose normative punishment value is bigger than the value of the importance of the goals; thus the agent feels pressured to comply with the norm to avoid punishments. In (Neto et al., 2011) the authors present the NBDI approach, which considers the normative contribution generated by evaluation between: (i) the norms' rewards and punishments, and (ii) the importance of the goals.

We chose these examples to compare with our approach because they represent the most common strategies followed by agents when they face a norm compliance decision. Our approach is based on (Neto et al., 2011) and was improved by adding personality traits. A normative conflict is identified when different norms are active and have opposite deontic concepts. The norm contribution is then evaluated for each one of the conflicting norms and there are a few steps to follow: (i) for each goal, its importance is increased by a weight given to a personality trait, (ii) for each goal allowed by a norm (the norm does not restrict this goal), the norm contribution is increased, adding the importance of the goal, and (iii) for each norm that is active at the same time and has opposite deontic concepts, 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 restricts is created,

(ii) for each non-conflicting norm, the norm contribution is increased adding the norm contribution value to each norm in this set that restricts 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 the observation of how both the social contribution and the agent's individual satisfaction change, according to the norm compliance strategy it chooses, 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 by the number of times the agent has fulfilled the norms addressed to it. The agent's individual satisfaction is the number of goals achieved in relation to the number of goals generated.

We reproduced the experiment created in (López and Márquez, 2004) using all of these different approaches and comparing them with our approach. 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 their importance. In addition, each goal might have a personality trait associated, meaning that 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 well as the deontic concept and activation time. Thus, the norms are evaluated by agents following different strategies so that similar inputs produce different outcomes.

We observed both the social contribution and the agent's individual satisfaction taking into account the different percentages of normative conflicts over a period of time. First, no conflicts were 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 conflicted among themselves. Each experiment consisted of 100 runs, and in each run, 10 goals and 10 norms were used.

Table 5 and Table 6 show the properties of random norms and random goals used in this experiment, respectively.

Table 5: Random Norm Properties.

Property	Description
Addressee	Agent "X"
Activation	All norms are activated
Expiration	There is no expiration
Rewards	Random value in the range [0,5]
Punishments	Random value in the range [0,5] + Set of goals restricted by this norm
Deontic Concept	Random value in the range [-1,1], where -1 represents a prohibition, 0 represents a permission and 1 represents an obligation

Table 6: Random Goal Properties.

Property	Description
Name	Goal + random value in the range [0,9]
Value	Random value in the range [0,5]
Norm Required	Random set of norms
Personality Trait	Two personality traits with a random value [5,10]
Belief Required	No belief was required

First, the Pressured strategy shows that the agent achieves more individual goals rather than contributes to the society. Figure 8 shows the agent's behavior in different conflicting norm situations.

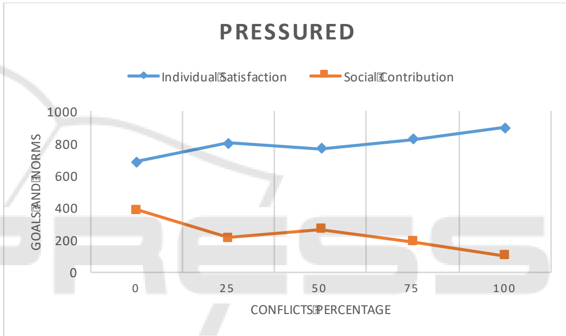

Figure 8: Pressured Strategy.

The Social strategy shows that initially, with no-conflicting norms, the agent fulfills all norms because, first, the agent complies with all the adopted active norms and then decides which goals will be achieved. Figure 9 shows that, as a result, the number of goals achieved increases gradually.

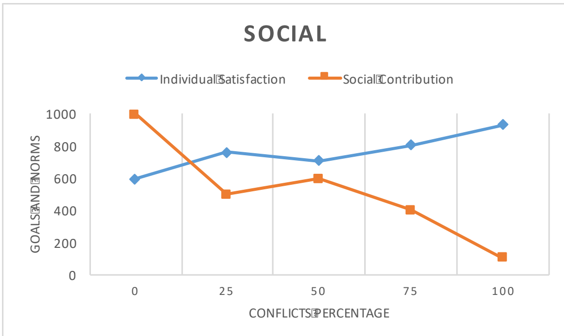

Figure 9: Social Strategy.

Agents using the Rebellious strategy violate all norms and, as no goal are restricted, all of them will be achieved. Figure 10 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 violating norms that restrict its goals.

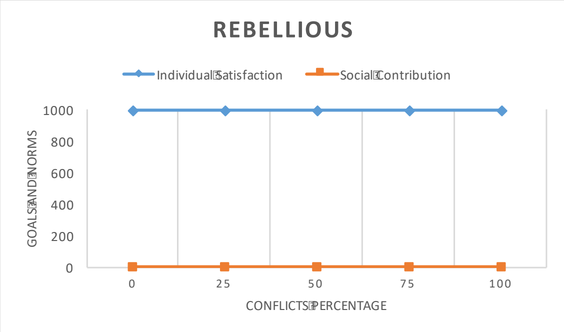

Figure 10: Rebellious Strategy.

The agent using the NBDI strategy considers the value of the social contribution to fulfill, or violate, each norm before deciding to comply with it, or not. Figure 11 shows that more goals are achieved when the normative conflicts increase.

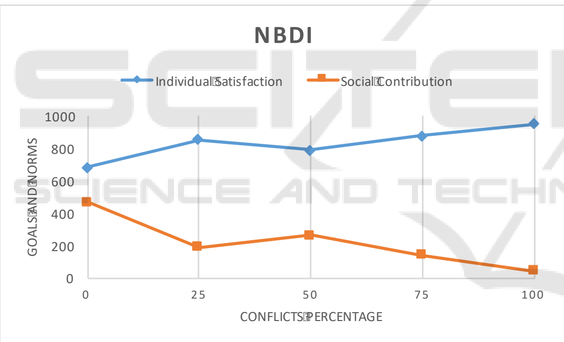

Figure 11: NBDI Strategy.

The Personality traits strategy considers the norm contribution developed in NBDI adding the personality traits value. The experiment results are similar to the NBDI strategy, although the agent meets more individual goals. Figure 12 shows the agent's behavior regarding norms compliance and goals achievement.

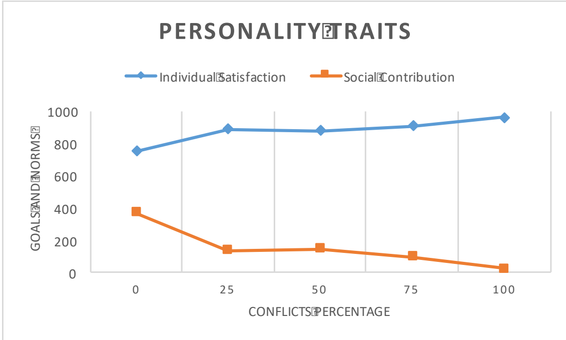

Figure 12: Personality Traits Strategy.

As can be observed, the personality traits strategy 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 the number of individual goals.

Figure 13 shows the comparison between all of the five strategies. As a result, the Personality Traits strategy achieved more goals than the Social, Pressured and NBDI strategies. It shows that the Personality Traits strategy helps the agent to fulfill more individual goals and increases the individual satisfaction.

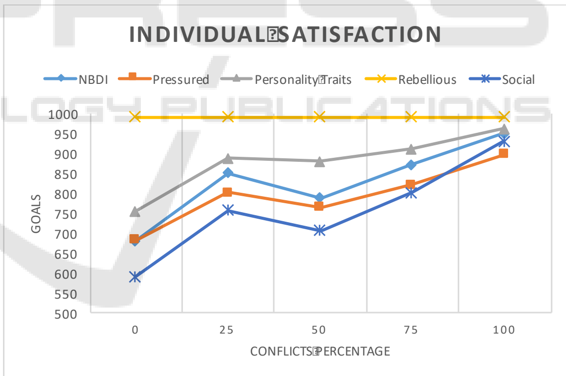

Figure 13: Individual satisfaction overview.

Figure 14 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 fulfills fewer norms than other strategies, except the Rebellious strategy, which always violates all the norms. Thus, the developed approach is a middle ground between Rebellious strategy and NBDI strategy.

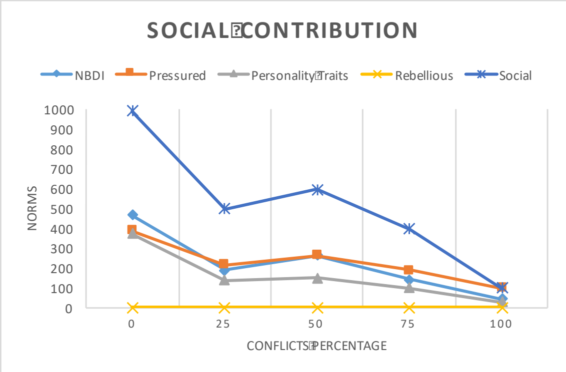

Figure 14: Social Contribution overview.

6 CONCLUSION AND FUTURE WORK

This paper proposes an approach to deal with normative conflicts 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 change 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 with different behaviors. The BDI-agent with personality traits 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, such 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 for the agent's current health state and how fast the recovery takes place 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 that take personality traits into account. For example: (i) in risk areas, where firefighters are responsible for planning people's evacuation, and (ii) in crime prevention, where the police are responsible for arresting criminals and keeping civilians safe. Last but not least, we will apply these different strategies to environments that have more agents, in order to analyze their behavior and evaluate the norms addressed to the agent, and the agent's internal goals.

REFERENCES

Barbosa, S. D. J., Silva, F. A. G. Da, Furtado, A. L., Casanova M. A., 2015. Plot Generation with Character-Based Decisions, In *Comput. Entertain. 12, 3*, Article 2.

Bordini, R. H.; Hübner, J. F.; Wooldridge, M., 2007. *Programming Multi-Agent Systems in AgentSpeak using Jason*. Wiley-Blackwell. 1st edition.

Bratman, M., 1987. *Intention, plans, and practical reason*, Harvard University Press.

Criado, N., Argente, E., Noriega, P., and Botti, V., 2010. Towards a Normative BDI Architecture for Norm Compliance, In *COIN@ MALLOW2010*, p. 65-81.

Jones, H., Saunier, J., Lourdeaux, D., 2009. Personality, Emotions and Physiology in a BDI Agent Architecture: The PEP -> BDI Model, In *IEEE*, p. 263-266.

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

López, L. F., Márquez, A. A., 2004. An Architeture for Autonomous Normative Agents, *IEEE*, Puebla, México.

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

Neto, B. F. S., Silva, V. T., Lucena, C. J. P., 2011. Developing Goal-Oriented Normative Agents: The NBDI Architecture, In *International Conference on Agents and Artificial Intelligence*.

Pereira, D., Oliveira, E., Moreira, N., Sarmento, L., 2005. Towards an Architecture for Emotional BDI Agents, In *Portuguese conference on artificial intelligence.*

Rao, A. S., Georgeff, M. P., 1995. *BDI agents: From theory to practice*, ICMAS. Vol. 95.

Vasconcelos, W.W., Kollingbaum, M.J., Norman, T. J., 2007. 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.

Viana, M. L., Alencar, P., Everton Guimarães, Cunha, F. J. P., 2015. JSAN: A Framework to Implement Normative Agents, In *The 27th International Conference on Software Engineering & Knowledge Engineering*, p. 660-665.

Viana, M. L., Alencar, P., Lucena, C. J. P., 2016. A Modeling Language for Adaptive Normative Agents. In *EUMAS, 2016, Valencia. European Conference on Multi-Agent Systems*.

Wooldridge, M. Ciancarini, P., 2001. Agent-Oriented Software Engineering: The State of the Art, In *LNCS 1957*.

Wooldridge, M., Jennings, N. R., Kinny, D., 1999. A methodology for agent-oriented analysis and design, In *Proceedings of the Third International Conference on Autonomous Agents Agents 99*.