A Probabilistic Implementation of Emotional BDI Agents

João Gluz and Patricia Jaques
PIPCA - UNISINOS, Av. Unisinos, 950, Bairro Cristo Rei, CEP 93.022-000, São Leopoldo, Brazil

Keywords: BDI, Emotions, Appraisal, OCC.

Abstract: A very well known reasoning model in Artificial Intelligence is the BDI (Belief-Desire-Intention). A BDI agent should be able to choose the more rational action to be done with bounded resources and incomplete knowledge in an acceptable time. Although humans need emotions in order to make immediate decisions with incomplete information, traditional BDI models do not take into account affective states of the agent. In this paper we present an implementation of the appraisal process of emotions in BDI agents using a BDI language that integrates logic and probabilistic reasoning. Specifically, we implement the event-generated emotions with consequences for self based on the OCC cognitive psychological theory of emotions. We also present an illustrative scenario and its implementation. One original aspect of this work is that we implement the emotions intensity using a probabilistic extension of a BDI language. This intensity is defined by the desirability central value, as pointed by the OCC model. In this way, our implementation of an emotional BDI allows to differentiate between emotions and affective reactions. This is an important aspect because emotions tend to generate stronger response. Besides, the intensity of the emotion also determines the intensity of an individual reaction.

1 INTRODUCTION

A very well known reasoning model in Artificial Intelligence (AI) is the BDI (Belief-Desire-Intention). It views the system as a rational agent having certain mental attitudes of belief, desire, intention, representing, respectively, the informational, motivational and deliberative states of the agent (Rao and Georgeff, 1995; Wooldridge, 2009). A rational agent has bounded resources, limited understanding and incomplete knowledge on what happens in the environment where it lives in. A BDI agent should be able to choose the more rational action to be done with bounded resources and incomplete knowledge in an acceptable time.

Damasio (Damasio, 1994) showed that humans need emotions in order to make immediate decisions with incomplete information. BDI agents also need to decide quickly and with incomplete data from the environment. The BDI is a practical reasoning architecture, that is reasoning directed towards action, employed in the cases when the environment is not fully observable (Russell and Norvig, 2010). However, most BDI models do not take into account the agent's emotional mental states in its process of decision making.

Among the several approaches of emotions - for example, basic emotions (Ekman, 1992), dimensional models (Plutchik, 1980), etc, the appraisal theory (Scherer, 2000; Scherer, 1999; Moors et al., 2013) appears to be the most appropriate to implement emotions into BDI agents. According to this theory, emotions are elicited by a cognitive process of evaluation called appraisal. The appraisal depends of one's goals and values. These goals and values can be represented as the BDI agents' goals and beliefs. In this way, it is possible to make a direct relation between BDI agents goals and beliefs and appraisals.

When addressing emotional mental states in BDI agents, several research questions should be addressed. An important issue is how to represent and implement the emotional appraisal, the emotions and its properties, such as, intensity. This is a first step before representing how emotions can interfere back in the cognitive processes of the agent, such as decision making. When addressing emotions, we should take into account the emotions intensity, since it defines when an emotion will occur or not. When an affective reaction does not achieve a sufficient intensity threshold, it will not be experienced as an emotion.

The formal logical BDI approach is not appropriate to represent the emotions intensity because it does not allow to represent imprecise data. This is the reason because most emotional BDI models do not take

Gluz J. and Jaques P..
A Probabilistic Implementation of Emotional BDI Agents.
DOI: 10.5220/0004815501210129
In *Proceedings of the 6th International Conference on Agents and Artificial Intelligence* (ICAART-2014), pages 121-129
ISBN: 978-989-758-015-4

into account the notion of emotion intensity (Jiang et al., 2007; Van Dyke Parunak et al., 2006; Adam et al., 2009). In order to represent the emotions intensity, we use AgentSpeak(PL) BDI language (Silva and Gluz, 2011). AgentSpeak(PL) is a new agent programming language, which integrates BDI and probabilistic reasoning, i. e., bayesian networks.

This article presents an implementation of the emotional appraisal into BDI agents. We are interested in the implementation of emotions and their intensity using a language that integrates BDI and Bayesian reasoning. Unlike other works, as we use a probabilistic extension of the BDI model that is able to represent the intensity of the affective reactions, our work is able to differentiate between emotions and other affective states with low intensity. Besides, the intensity of the emotion also determines how strong is the response of an individual (Scherer, 2000).

In this paper, we focus on the event-generated emotions, i. e., emotions that are elicited by the evaluation of the consequences of an event for the accomplishment of a person's goals. We do not formalize emotions in which their appraisal evaluate the consequences for others, such as resentment, pity, gloating and happy-for. We chose to implement the event-generated emotions with consequences for self, since these emotions seems to be the most important in the decision making process (Bagozzi et al., 2003; Isen and Patrick, 1983; Raghunathan and Pham, 1999).

This paper is organized as follow. Section 2 presents the OCC model, the psychological emotional model that grounds our work. In Section 3, we describe the AgentSpeak(PL), a language that integrates BDI and Bayesian Decision Network to reason about imprecise data. Section 4 compares the proposed work with related works and highlights its main contribution. In Section 5, we cite a scenario and its implementation with AgentSpeak (PL) to illustrate how it can be used to implement emotional probabilistic BDI agents. Finally, in Section 6, we present some conclusions.

2 THE OCC MODEL

This work proposes an extension of the BDI model to integrate emotions. We intend to implement the appraisal process of emotions into BDI agents. We implemented the appraisal process according to the OCC model (Ortony et al., 1990) and using probabilistic reasoning to represent emotions intensity.

According to the cognitive view of emotions (Scherer, 1999), emotions appear as a result of an evaluation process called appraisal. The central idea of the appraisal theory is that "the emotions are elicited and differentiated on the basis of a person's subjective evaluation (or appraisal) of the personal significance of a situation, event or object on a number of dimensions or criteria" (Scherer, 1999).

Ortony, Clore and Collins (Ortony et al., 1990) constructed a cognitive model of emotion, called OCC, which explains the origins of 22 emotions by describing the appraisal of each one. For example, hope appears when a person develops an expectation that some good event will happen in the future.

The OCC model assumes that emotions can arise by the evaluation of three aspects of the world: events, agents, or objects. **Events** are the way that people perceive things that happen. **Agents** can be people, biological animals, inanimate objects or abstractions such as institutions. **Objects** are objects viewed qua objects. There are three kinds of value structures underlying perceptions of goodness and badness: goals, standards, and attitudes. The events are evaluated in terms of their desirability, if they promote or thwart one's goals and preferences. Standards are used to evaluate actions of an agent according to their obedience to social, moral, or behavioural standards or norms. Finally, the objects are evaluated as appealing depending on the compatibility of their attributes with one's tastes and attitudes. In this paper we refer to the emotions that are generated from the evaluation of an event consequences according to one's goals as **event-generated emotions**.

The elicitation of an emotion depends on a person's perception of the world – his construal. If an emotion such as distress is a reaction to some undesirable event, the event must be construed as undesirable. For example, when one observes the reactions of players at the outcome of an important game, it is clear that those on the winning team are elated while those on the losing team are devastated. In a real sense, both the winners and losers are reacting to the same objective event. It is their construal of the event that is different. The winners construe it as desirable, while the others construe it as undesirable. It is this construal that drives the emotion system.

A central idea of the model is the type of an emotion. An emotion type is a distinct kind of emotion that can be realized in a variety of recognizably related forms and which are differentiated by their intensity. For example, fear is an emotion type that can be manifested in varying degrees of intensity, such as "concern" (less afraid), "frightened", and "petrified" (more afraid). The use of emotion type has the goal of being language-neutral so that the theory is universal, independent of culture. Instead of defining an emotion by using English words (the author's language), the emotions are characterized by their eliciting conditions.

In the OCC model, the emotions are also grouped according to their eliciting conditions. For example, the "attribution group" contains four emotion types, each of which depends on whether the attribution of responsibility to some agent for some action is positive or negative, and on whether the agent is the self or another person.

The OCC model is illustrated in Figure 1. When goals are the source, one may feel pleased if the event is desirable, or displeased if it is not. Which specific emotion arises depends on whether the consequences are for other or for oneself. When concerned for oneself (label as CONSEQUENCES FOR SELF), the evaluation depends on whether the outcomes are past (label as PROSPECTS IRRELEVANT), like joy and distress, or prospective (label as PROSPECTS RELEVANT), such as hope and fear. If the prospect is confirmed or not, other four emotions may arise, such as satisfaction, disappointment, fear-confirmed and relief. When concern for other (label as CONSEQUENCES FOR OTHER), the outcomes are evaluated according to when they are undesirable (label as UNDESIRABLE FOR OTHER), such as gloating and pity, or desirable for other (label as DESIRABLE FOR OTHER), such as happy-for, resentment.

When the actions of agents are evaluated according to standards, affective reactions of approval or disapproval arise. The specific emotions depend on whether the action is one's own (labelled as SELF AGENT), such as pride and shame; or someone else's (labelled as OTHER AGENT), such as admiration and reproach.

The aspects of an object are evaluated according to one's tastes, if one likes or dislikes. In this case, emotions such as love and hate may arise. Finally, emotions like anger and gratitude involve a joint focus on both goals and standards at the same time. For example, one's level of anger depends on how undesirable the outcomes of events are and how blameworthy the related actions are.

According to the OCC model, affective reactions are effectively experienced as emotions only if they achieve a minimum intensity degree. Before this, these affective reactions have only a potential for the emotion. But, after this potential surpasses the minimum threshold necessary for an emotion, the emotion starts to be felt.

The OCC authors believe that this model when implemented in a machine can help to understand what emotions people experience under what conditions. According to them, it is not the objective of the OCC model to implement machine with emotions,

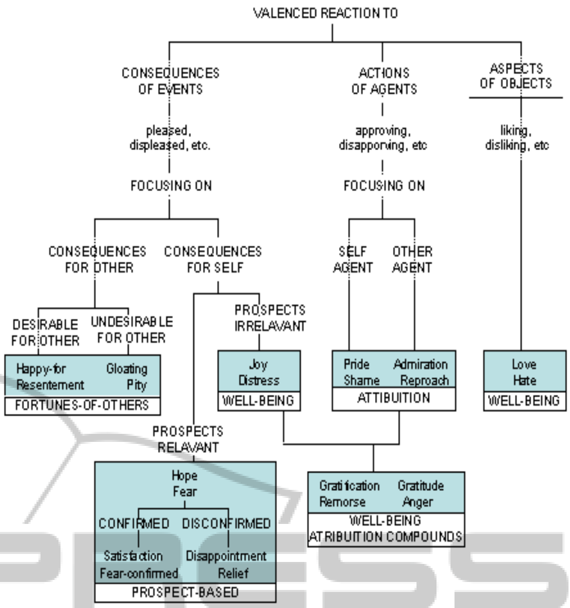

Figure 1: Global Structure of Emotion Types - OCC Model (Ortony et al., 1990).

but to be able to predict and explain human cognitions related to emotions. However, Picard (Picard, 2000) disagrees and believes that the OCC model can be used for emotion synthesis in machines. In fact, in computing science research, there already are several works that use OCC in order to implement emotions in machine (Gebhard, 2005; Dias and Paiva, 2013; Jaques et al., 2011; Signoretti et al., 2011).

The OCC authors also agree that this model is a highly oversimplified vision of human's emotions, since in reality a person is likely to experience a mixture of emotions (Ortony et al., 1990). Otherwise, in order to understand which set of emotions is a person or agent experiencing, we must first try to identify each emotion separately. This is the case of the approach adopted in the proposed work and in great part of the emotion synthesis research.

3 A HYBRID BDN+BDI MODEL

The BDI model was based on the works of Searle and Dennet, posteriorly generalized by the philosopher Michael Bratman (Bratman, 1990), which gave particular attention to the role of intentions in reasoning (Wooldridge, 1999). The BDI approach views the system as a rational agent having certain mental attitudes of belief, desire, intention, representing, respectively, the information, motivational and deliberative states of the agent (Rao and Georgeff, 1995). A rational agent has bounded resources, limited understanding and incomplete knowledge on what happens in the environment it lives in.

The **beliefs** represent the information about the state of the environment that is updated appropriately after each sensing action. **Desires** are the motivational state of the system. They have information about the objectives to be accomplished, i. e. what priorities or pay-offs are associated with the various current objectives. They represent a situation that the agent wants to achieve. The fact that the agent has a desire does not mean that the agent will do it. The agent carries out a deliberative process in which it confronts its desires and beliefs and chooses a set of desires that can be satisfied. An **intention** is a desire that was chosen to be executed by a plan, because it can be carried out according to the agents beliefs (it is not rational an agent carries out something that it does not believe). Plans are pre-compiled procedures that depend on a set of conditions for being applicable. The desires can be contradictory to each other, but the intentions cannot (Wooldridge, 1999). Intentions represent the goals of the agent, defining the chosen course of action. An agent will not give up on its intentions – they will persist, until the agent believes it has successfully achieved them, it believes it cannot achieve them or because the purpose of the intention is no longer present.

In the BDI model, a belief is defined as a two-state logical proposition: or the agent believes that a certain event is true or it believes that the event is false. Today, programming languages and tools available for development BDI agents do not work with the concept of probabilistic beliefs (Bordini et al., 2005), i.e. they do not allow agents to understand, infer or represent degrees of belief (or degrees of uncertainty) about a given proposition. A degree of belief is defined by the subjective probability assigned to a particular belief.

The concept of Bayesian Networks (BN) (Pearl, 1988) fits in this scenario, allowing to model the probabilistic beliefs of some agent. BN alone are excellent tools to represent probabilistic models of agents, but, with the addition of utility, and decision nodes, it is possible to use the full spectrum of Decision Theory to model agent's behaviour. BN extended with utility, and decision nodes, are called Bayesian Decision Networks (BDN) (Russell and Norvig, 2010).

The integration between the current agent programming languages and the concept of belief probabilities can be approached in several ways and at different levels of abstraction. In practical terms, it is possible to make an *ad-hoc* junction of both kind of models in the actual programming code of the agents. Both BDI and BDN agent programming environments rely on libraries and development frameworks, with a standard Application Programming Interface (API). Thus, a hybrid agent can be designed and implemented by combining calls from different sets of APIs, each one from distinct programming environments. The more abstract level to address this issue of integration is usually treated by Probabilistic Logics (Korb and Nicholson, 2003). Although Probabilistic Logics have the ability to represent both logical beliefs and probabilistic beliefs, there are notorious problems related to the tractability of the resulting models (Korb and Nicholson, 2003). Another approach is to extend logic programming languages (essentially Prolog) to handle probabilistic concepts. P-Log (Baral and Hunsaker, 2007) and PEL (Milch and Koller, 2000) fit in this category. They offer interesting ideas, but they lack the full integration with BDI programming languages, like AgentSpeak(L) (Bordini et al., 2007).

In the present work we consider a programming approach that fully integrates the theoretical and practical aspects of BDN and BDI models. We use AgentSpeak(PL) (Silva and Gluz, 2011) to implement a probabilistic process, for the appraisal of (some of) the OCC's emotions. AgentSpeak(PL) is a new agent programming language, which is able to support BDN representation and inference in a seamless integrated model of beliefs and plans of an agent. AgentSpeak(PL) is based on the language AgentSpeak(L) (Bordini et al., 2007), inheriting from it all BDI programming concepts. AgentSpeak(PL) is supported by JasonBayes (Silva and Gluz, 2011), an extension of the Jason (Bordini et al., 2007) agent development environment. The main changes of AgentSpeak(PL), in respect to AgentSpeak(L), were:

- Inclusion of a probabilistic decision model of the agent, consisting of the specification of a BDN.
- Inclusion of events/triggers based on probabilistic beliefs.
- Inclusion of achievement and test goals, based on probabilistic beliefs.
- Inclusion of actions, which are able to update the probabilistic model.

4 RELATED WORK

There are other researchers that have also been working in the extension of the BDI architecture in order to incorporate emotions. (Jiang et al., 2007) defines an extension of the generic architecture of BDI agents that introduces emotions. However, this work explores the possible influences of emotions to determine agents' beliefs and intentions. (Van Dyke Parunak et al., 2006) propose an extension of the BDI model to integrate emotions, which is based on the OCC model. In this model, the agent's beliefs interfere in its appraisal, which will determine its emotions. The emotions, on the other hand, interfere in the choice of intentions. This work is focussed on the impact of the emotions in agent's choice of intentions from desires. Other related works are the ones proposed by (Adam et al., 2009) and (Steunebrink et al., 2012; Steunebrink et al., 2008). Both works present a purely logical formalization of the OCC model. They are theoretical works that use a BDI modal logic to describe the OCC's appraisal. They do not show how to create an operational computing model based on the logical formalization.

All the above cited related works do not use a probabilistic model to represent the intensity of the emotions. Their model is not able to represent if the potential of an affective reaction in the appraisal process achieved the necessary threshold to elicit an emotion. In these works, when the appraisal process occurs, an agent always has an emotion. This is not the case, because an emotion only occurs when the intensity achieves a specific threshold (Ortony et al., 1990). To differentiate emotions from other affective reactions is important because emotions tend to generate stronger response. Besides, the intensity of the emotion also determines how strong is the response of an individual (Scherer, 2000).

5 A SCENARIO

Let us present a scenario in order to illustrate the process of appraisal that happens with an agent, which has an emotional BDI+BDN architecture. Let us consider a vacuum cleaner robot example, with its environment of a grid with two cells. Besides the desires to work and clean the environment, our agent, whose name is Vicky, also has the desire of protecting its own existence with a stronger priority.

While Vicky is cleaning Cell B, Nick, a clumsy researcher who also works at the same laboratory, goes towards Vicky without noticing it. When Vicky perceives that Nick is going to stomp on itself, the desire of self-protection becomes an intention and Vicky feels fear of being damaged.

In order to alert the awkward scientist, Vicky emits an audible alarm. Nick perceives Vicky (almost behind his feet) and step on the other cell, avoiding trample over Vicky. When Vicky perceives that it is not in danger any more, finally, it feels relief. It can continue to do its work; at least while Nick does not decide to come back to his office.

In the next section we present how the process of emotion appraisal, with the corresponding behaviour consequences, can be implemented in AgentSpeak(PL), and how this process will evolve in the the first reasoning cycles of the agent.

5.1 AgentSpeak(PL) Implementation

Emotions in the OCC model depend on several cognitive *variables*, related to the mental state of the agent. These variables are instrumental in the process of emotion's arousal, because they determine the intensity of the emotions. Not all affective reactions (evaluation of an event as un/desirable) are necessarily emotions: "Whether or not these affective reactions are experienced as emotions depends upon how intense they are" (Ortony et al., 1990, p. 20).

The OCC model divides these variables in central variables, like *desirability*, and local variables, like *effort*. For our formal model, we need to define the desirability and undesirability central variables because these variables are important to event-generated emotions.

Desirabilities are not utilities. From an ontological perspective they are based on distinct things: the prospective gains of some event in respect to the subjective value of agent goals versus the functional representation of rational preferences for states of the world. However, from a purely formal point of view, it is possible to use utilities to estimate the value of the desirability. If we consider a subjective model for the state of the world, then it is possible to assume that utilities (as a kind of evaluation) could enter in the formation of the desirability. The desirability of some event could be calculated as the difference between the utility of the state of the world previous to the occurrence of the event, and the utility after this event happens, as hypothesized by the agent.

Using this interpretation, it is possible to apply a tool like a Bayesian Decision Network (BDN) (Russell and Norvig, 2010) to make subjective (bayesian) probabilistic model of the states of the world, and to associate utilities to these states to estimate the desirability. The BDN diagram shown in Figure 2 represents Vicky's probabilistic model about the desirability of smashed and dirty events.

The tables presented in Table 1 define the prior and conditional probabilities, as well as the utility of this probabilistic model. These tables provide estimations of how good is to work and to continue to exist. They can also be used to estimate how much the fact of finding dirt will advance cleaning and self-

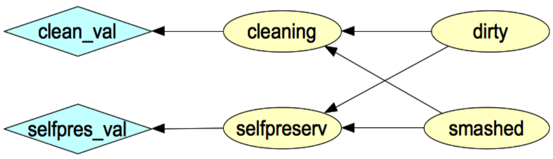

Figure 2: A BDN representing Vicky's probabilistic model about desirability of events.

Table 1: Probability tables for Vicky's BDN model.

smash	*P(smash)*	*dirty*	*P(dirty)*
true	0.5	true	0.5
false	0.5	false	0.5

cleaning	*smash*	*dirty*	*P(cleaning | dirty & smash)*
true	true	true	0.0
true	true	false	0.0
true	false	true	0.9
true	false	false	0.2
false	true	true	1.0
false	true	false	1.0
false	false	true	0.1
false	false	false	0.8

selfpreserv	*smash*	*dirty*	*P(selfpreserv | dirty & smash)*
true	true	true	0.1
true	true	false	0.1
true	false	true	1.0
true	false	false	1.0
false	true	true	0.9
false	true	false	0.9
false	false	true	0.0
false	false	false	0.0

cleaning	*U(clean_val)*
true	80
false	0

selfpreserv	*U(selfpres_val)*
true	100
false	0

preservation goals. Together with the knowledge that any smashing will severely hinder the agent's goals, this model will allow Vicky to estimate the desirability of cleaning and if it feels joy or fear.

Now, using an appropriate agent programming language, it is a relatively straightforward task to program Vicky's beliefs, emotions, intentions, and plans. To do so, we will use AgentSpeak(PL) (Silva and Gluz, 2011), which generalizes AgentSpeak(L), allowing a seamless integration of bayesian decision model in the BDI planning process.

An AgentSpeak(PL) source code contains a BDN model, the non-probabilistic beliefs, the agent's goals, and its corresponding plans. Figure 3 shows the source code[1] for the BDN presented in Figure 2.

```
// Probabilistic model
// Prior probabilities - standard syntax
%dirty(true) = 0.5.
%dirty(false) = 0.5.
// Prior probabilities - compact syntax
%smash(true,false) = [0.5, 0.5].
// Conditional probabilities - standard syntax
%cleaning(true) | smash(true) & dirty(true) = 0.0 .
%cleaning(true) | smash(true) & dirty(false) = 0.0 .
%cleaning(true) | smash(false) & dirty(true) = 0.9 .
%cleaning(true) | smash(false) & dirty(false) = 0.2.
%cleaning(false) | smash(true) & dirty(true) = 1.0.
%cleaning(false) | smash(true) & dirty(false) = 1.0.
%cleaning(false) | smash(false) & dirty(true) = 0.1.
%cleaning(false) | smash(false) & dirty(false)= 0.8.
// Conditional probabilities - compact syntax
%selfpreserving(true,false) | smash & dirty =
              [0.1, 0.1, 1.0, 1.0, 0.9, 0.9, 0.0, 0.0].
// Utility function - standard syntax
$clean_val | cleaning(true) = 80.0 .
$clean_val | cleaning(false) = 0.0 .
// Utility function - compact syntax
$selfpres_val | selfpreserv = [100.0, 0.0].
```

Figure 3: Vicky's BDN model programmed in AgentSpeak(PL).

```
min_joy(50). min_distress(-100).
located(cell_b).
!cleaning.
!selfpreserv.
+dirty <-
    +perceived(dirty); +!cleaning.
+human_fast_approx <-
    +prospect(smash); +selfpreserv.
+human_fast_depart <-
    -prospect(smash).
```

Figure 4: Vicky's initial beliefs, goals, and perception plans.

Vicky's initial non-probabilistic beliefs, its primary goals, and the plans to handle its perceptions are defined in Figure 4. Initially, as the battery energy level of Vicky is high and it is located in a dirty cell (cell B), Vicky selects the goal "clean the cell" as an intention. We can also expect that Vicky has also a basic self-preservation instinct, i. e., it has the intention of preserving its own existence. The perception plans start intentions about what to do if some event is perceived, incorporating the common sense knowledge of Vicky about what happens when dirt is detected or when some human is approximating fast.

The planning knowledge of Vicky, in respect to what to do in its operation cycles, is relatively simple and can be specified by the set of AgeentSpeak(PL) plans defined in Figure 5. The only small issue is that in AgentSpeak(PL) it is necessary to start goals and plans to calculate the desirability of events. This is

[1]This code can be programmed by hand, or it can be automatically generated from a graphical model similar to the BDN presented in Figure 2 (see (Silva and Gluz, 2011) for details).

```
+!cleaning : perceived(dirty) ->
    +dirty(true);
    !joy_inten(dirty,JI); JI > 0;
    !fear_inten(smash,FI); FI = 0;
    .cleaning_action.
+!selfpreserv : expect(smash) <-
    !fear_inten(smash,FI);FI > 0;
    .alarm_action.
```

Figure 5: Vicky's plans for cleaning, and self preservation intentions.

due to how the marginal probabilities and the utility values of the BDN model are calculated: whenever some evidence is added to the beliefs of the agent, the bayesian inference engine is activated and all probabilities and utilities are recalculated. An evidence is simply a non-probabilistic belief, like `dirty(true)` or `smashed(false)`, that are identical in name to some variable (node) of the BDN model.

Vicky's code is completed by the plans to detect if it feels some joy, or fear, presented in the Figure 6. These plans implement the rule-based computational model for these emotions, as presented in (Ortony et al., 1990, p. 182-186), defining operational functions to estimate the potential and the intensity of these emotions.

5.2 First Reasoning Cycle

Initially, as the battery energy level of Vicky is high and it is located in a dirty cell (cell B), Vicky selects the goal "clean the cell" as an intention. We can also expect that Vicky has also a basic self-preservation instinct, i. e., it has the intention of preserving its own existence.

In the first cycle, when dirt is found, but not smashing event expected, then the variable `U1` in `joy_poten(dirty,JP)` plan will give the utility if it is assumed that dirt is found. The value of `U1` will be `172`, with the utility function returning: `$(clean_val) = 72`, and `$(selfpres_val)=100`. The variable `U2` will give the utility, if we assume that dirt is not found. Its value is `116`, with `$(clean_val) = 16`, and `$(selfpres_val)=100`. The difference `JP` will be `56`. With this value for the potential of joy, then `joy_inten(dirty,JI)` plan will return `JI=6`, indicating that Vicky is feeling joy.

5.3 Second Reasoning Cycle

In the second cycle, while it is cleaning, Vicky receives from its vision sensor the perception "Human fast approximation". Vicky also knows that human approximation can damage a robot, because the human can trample the robot. It learned that from a previous bad-succeed experience with Nick.

```
+!joy_poten(dirty,JP): not dirty(true) <-
    J = 0.
+!joy_poten(dirty,JP): dirty(true) <-
    U1 = $(clean_val) + $(selfpres_val);
    -+dirty(true);
    U2 = $(clean_val) + $(selfpres_val);
    -+dirty(true);
    JP = U1 - U2.
+!joy_inten(dirty,JI) <-
    !joy_poten(dirty,JP); ?min_joy(MJ);
    if (JP>MJ) {JI = JP - MJ} else {JI = 0}.

+!fear_poten(smash,J): not expected(smash) <-
    J = 0.
+!fear_poten(smash,J): prospect(smash) <-
    +smash(true),
    U1 = $(clean_val) + $(selfpres_val);
    -+smash(false);
    U2 = $(clean_val) + $(selfpres_val);
    -smash(false);
    F = U2 - U1.
+!fear_inten(dirty,FI) <-
    !fear_poten(smash,FP); ?min_fear(MF);
    if (FP>MF) {FI = FP - MF} else {FI = 0}.
```

Figure 6: Vicky's plans to estimate its joy, and fear emotions.

In the second cycle, a similar situation occurs, but this time the variable `U1` in the `fear_poten(smash,FP)` plan, which estimates the utility of some smashing occurs, will be `10`, with `$(clean_val) = 0`, and `$(selfpres_val)=10`. The value of `U2` in the same plan will be `172`, giving an estimation of `162` for `FP`, the undesirability of being smashed. Then, `fear_inten(dirty,FI)` plan will return `FI=62`, indicating that Vicky has fear.

6 CONCLUSIONS

In this paper we presented an implementation of the appraisal process of emotions in BDI agents using a BDI language that integrates logic and probabilistic reasoning. Specifically, we implemented the event-generated emotions with consequences for self based on the OCC cognitive psychological theory of emotions. We also presented an illustrative scenario and its implementation.

One original aspect of this work is that we implemented the emotions intensity using a probabilistic extension of a BDI language, called AgentSpeak(PL). This intensity is defined by the desirability central value, as pointed by the OCC model. In this way, our implementation of emotional BDI allows to differentiate between emotions and affective reactions. This is an important aspect because emotions tend to generate stronger response. Besides, the intensity of the emotion also determines how strong is the response of an individual (Scherer, 2000).

We are aware that the implementation of the appraisal of emotions is only a first step. An emotional BDI implementation should address other important dynamic processes between emotions and the mental states of desires, intentions and beliefs in the BDI architecture. As the BDI is a practical reasoning architecture, that is reasoning towards action (Wooldridge, 1999), it is important to discuss how the use of emotions can help the agent to choose the most rational action to be done and how the emotions can improve the way that an agent reasons or decides or acts. These are open questions that we intend to address in a future work. However, we believe that the implementation of the appraisal and the arousal of an emotion depending on the intensity of the affective reaction, presented in this paper, is an important and initial point since the appraisal evaluation explains the origin of an emotion and also differentiates them (Scherer, 1999).

ACKNOWLEDGEMENTS

This work is supported by the following research funding agencies of Brazil: CAPES, CNPq, FAPERGS and CTIC/RNP.

REFERENCES

Adam, C., Herzig, A., and Longin, D. (2009). A logical formalization of the OCC theory of emotions. *Synthese*, 168(2):201–248.

Bagozzi, R. P., Dholakia, U. M., and Basuroy, S. (2003). How effortful decisions get enacted: the motivating role of decision processes, desires, and anticipated emotions. *Journal of Behavioral Decision Making*, 16(4):273–295.

Baral, C. and Hunsaker, M. (2007). Using the probabilistic logic programming language p-log for causal and counterfactual reasoning and non-naive conditioning. In *Proceedings of the 20th international joint conference on Artifical intelligence*, IJCAI'07, pages 243–249, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Bordini, R. H., Dastani, M., Dix, J., and Seghrouchni, A. E. F. (2005). *Multi-agent programming : languages, platforms and applications.*, volume 15 of *Multiagent systems, artificial societies, and simulated organizations.* Springer, New York.

Bordini, R. H., Hübner, J. F., and Wooldridge, M. (2007). *Programming Multi-Agent Systems in AgentSpeak using Jason (Wiley Series in Agent Technology)*. John Wiley & Sons.

Bratman, M. (1990). What is intention? In Cohen, P. R., Morgan, J. L., and Pollack, M. E., editors, *Intentions in Communications*, Bradford books, pages 15–31. MIT Press, Cambridge.

Damasio, A. R. (1994). *Descartes' error : emotion, reason, and the human brain*. G.P. Putnam New York.

Dias, J. a. and Paiva, A. (2013). I want to be your friend: establishing relations with emotionally intelligent agents. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*, AAMAS '13, pages 777–784, Richland, SC. IFMAS.

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, 6(3-4):169–200.

Gebhard, P. (2005). Alma: a layered model of affect. In *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, AAMAS '05, pages 29–36, New York, NY, USA. ACM.

Isen, A. M. and Patrick, R. (1983). The effect of positive feelings on risk taking: When the chips are down. *Organizational Behavior and Human Performance*, 31(2):194–202.

Jaques, P. A., Vicari, R., Pesty, S., and Martin, J.-C. (2011). Evaluating a Cognitive-Based Affective Student Model. In D'Mello, S. K., Graesser, A. C., Schuller, B., and Martin, J.-C., editors, *International Conference on Affective Computing and Intelligent Interaction (ACII)*, volume 6974 of *Lecture Notes in Computer Science*, pages 599–608. Springer.

Jiang, H., Vidal, J., and Huhns, M. N. (2007). *International Conference On Autonomous Agents*. ACM, New York.

Korb, K. and Nicholson, A. (2003). *Bayesian Artificial Intelligence*. Chapman & Hall/CRC Computer Science & Data Analysis. Taylor & Francis.

Milch, B. and Koller, D. (2000). Probabilistic models for agents' beliefs and decisions. In *Proceedings of the Sixteenth conference on Uncertainty in artificial intelligence*, UAI'00, pages 389–396, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Moors, A., Ellsworth, P. C., Scherer, K. R., and Frijda, N. H. (2013). Appraisal Theories of Emotion: State of the Art and Future Development. *Emotion Review*, 5(2):119–124.

Ortony, A., Clore, G. L., and Collins, A. (1990). *The Cognitive Structure of Emotions*. Cambridge University Press.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

Picard, R. W. (2000). *Affective Computing*. University Press Group Limited.

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. *Emotion: Theory, research, and experience*, 1(3):3–33.

Raghunathan, R. and Pham, M. T. (1999). All negative moods are not equal: Motivational influences of anxiety and sadness on decision making. *Organizational Behavior and Human Decision Processes*, 79(1):56–77.

Rao, A. S. and Georgeff, M. (1995). BDI Agents: from Theory to Practice. Technical Report Technical Note 56, Melbourne, Australia.

Russell, S. J. and Norvig, P. (2010). *Artificial Intelligence: A Modern Approach*. Prentice Hall Series in Artificial Intelligence. Pearson Education/Prentice Hall.

Scherer, K. R. (1999). Appraisal theory. In Dalgleish, T. and Power, M., editors, *Handbook of Cognition and Emotion*, volume 19, chapter 30, pages 637–663. John Wiley & Sons Ltd.

Scherer, K. R. (2000). Psychological models of emotion. In Borod, J., editor, *The neuropsychology of emotion*, volume 137 of *The neuropsychology of emotion*, chapter 6, pages 137–162. Oxford University Press.

Signoretti, A., Feitosa, A., Campos, A. M., Canuto, A. M., Xavier-Junior, J. C., and Fialho, S. V. (2011). Using an affective attention focus for improving the reasoning process and behavior of intelligent agents. In *Proceedings of the 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology - Volume 02*, WI-IAT '11, pages 97–100, Washington, DC, USA. IEEE Computer Society.

Silva, D. and Gluz, J. (2011). AgentSpeak(PL): A New Programming Language for BDI Agents with Integrated Bayesian Network Model. In *2011 International Conference on Information Science and Applications*. IEEE.

Steunebrink, B., Dastani, M., and Meyer, J.-J. (2012). A formal model of emotion triggers: an approach for bdi agents. *Synthese*, 185(1):83–129.

Steunebrink, B. R., Meyer, J.-J. C., and Dastani, M. (2008). A Formal Model of Emotions: Integrating Qualitative and Quantitative Aspects. In Bordini, R., Dastani, M., Dix, J., and Fallah-Seghrouchni, A. E., editors, *ECAI 2008*. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany.

Van Dyke Parunak, H., Bisson, R., Brueckner, S., Matthews, R., and Sauter, J. (2006). A model of emotions for situated agents. *Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems AAMAS 06*, 2006:993.

Wooldridge, M. (1999). Intelligent Agents. In Weiss, G., editor, *Multiagent systems*, pages 27–77. MIT Press, Cambridge, MA, USA.

Wooldridge, M. (2009). *An Introduction to MultiAgent Systems*. John Wiley & Sons.