The Knowledge Engineering Review, Vol. 34, e26, 1 of 33.
doi:10.1017/S0269888919000122

Designing emotional BDI agents: good practices and open questions

YANET SÁNCHEZ-LÓPEZ and EVA CEREZO

Department of Computer Science and Systems Engineering, Aragón Institute of Engineering Research, University of Zaragoza, Zaragoza, Spain
e-mails: ysanchezl@unizar.es; ecerezo@unizar.es

Abstract

Intelligent agents built on the basis of the BDI (belief–desire–intention) architecture are known as BDI agents. Currently, due to the increasing importance given to the affective capacities, they have evolved giving way to the BDI emotional agents. These agents are generally characterized by affective states such as emotions, mood or personality but sometimes also by affective capacities such as empathy or emotional regulation. In the paper, a review of the most relevant proposals to include emotional aspects in the design of BDI agents is presented. Both BDI formalizations and BDI architecture extensions are covered. From the review, common findings and good practices modeling affect, empathy and regulatory capacities in BDI agents, are extracted. In spite of the great advance in the area several, open questions remain and are also discussed in the paper.

1 Introduction

The definition of intelligent agent is a controversial issue, but it can be said that is an autonomous entity that receives dynamic surrounding information through sensors and acts on the environment through actuators, showing a goal-directed behavior (Russell *et al.*, 2003). The modeling of the cognitive processes into intelligent agents is based on different theories (Moore, 1980; Bratman, 1987; Newell, 1994) that explain, from different points of view, the functioning of the human mind. Intelligent agents implemented over the base of a cognitive theory are known as cognitive agents. The most developed ones are those based on cognitive architectures, such as Soar (Laird *et al.*, 1987), ACT-R (Anderson, 1993) and BDI (belief–desire–intention) (Rao & Georgeff, 1995). BDI architecture, which exists with several variants, is one of the most used software architectures for intelligent agents because of its simplicity and robustness. It was born on the basis of the theoretical foundation of Michael Bratman's practical reasoning (Bratman, 1987) and it is based on three fundamental mental attitudes: beliefs (that represent the information about the environment and oneself), desires (that represent the motivational state of the agent) and intentions (actions plans that have been selected and are committed to achieve and that provide the deliberative character to the model). Intelligent agents built on the basis of the BDI architecture are known as BDI agents and are the scope of this paper.

The inclusion of emotional aspects in the design of intelligent agents is become central issue in agent's research (Hudlicka, 2011). Firstly, because emotions can enhance the agent's autonomy and effectiveness, particularly in complex and uncertain environments, by biasing the agent's behavior toward a selection of particular goals and actions, producing a more adaptive behavior. Secondly, because emotions can enhance the agent's believability, awarding affective and social realism to the agent. Not only will the agent exhibit more realistic behavior (showing anger, tiredness or relief depending on the situation) but will be capable of reacting differently to the same situation or event depending on changing moods and goals. Emotions may also play a significant role in social and interpersonal domains, as different types of emotions could be exhibited depending on the relationship with the user or other agents.

This is why emotions may significantly raise agent's effectiveness when interacting with a human and are extensively used in interface agents and, more specifically, in embodied agents. Agent's embodiment is closely related to emotional modeling as it conditions the expressive channels: facial expressions, gaze, body postures and speech. Agent's body can be either virtual or physical (Serenko *et al.*, 2007). An example of a physically embodied agent is Kismet, a sociable robot that shows emotions and expressive behavior in social interaction with humans (Breazeal, 2003) and Yuppie, an emotional pet robot capable of interacting with humans and show affective behavior (Velásquez & Maes, 1997). In the case of virtual embodied agents, they can be found many applications: in medical domains (Bickmore *et al.*, 2016), in virtual storytelling (Gris *et al.*, 2016), as museum guides (Kopp *et al.*, 2005), as interviewers (Nunamaker *et al.*, 2011) or in virtual role-playing (Emonts *et al.*, 2012). In these and all other applications in which credible human-like behaviors are pursued the consideration not only of cognitive but affective capabilities is a must. Nevertheless, the inclusion of affective aspects in the modeling of agents is not trivial and different approaches can be found in the literature as it will be discussed in the next section.

2 Incorporating affect into intelligent affective agents

The design of intelligent agents has focused in the last years on the modeling of emotions and their influence on the cognitive process, pursuing a more credible and human-like behavior. The modeling of emotions and their influence on the cognitive process is not the only aspect that has attracted the attention of developers but also the modeling and influence of other affective states such as mood, personality and other affective capacities like empathy, emotional regulation and coping.

The works aimed to incorporate human-like emotions mechanisms into intelligent agents can be grouped into three categories (following the classification of computational models of emotions presented in Reisenzein *et al.* (2013)): formalizations, extensions of existing cognitive and agents architectures and specialized affective architectures.

Formalizations consists of translating a psychological emotion theory (or selected assumptions of one or more emotion theories), as an intermediate step, into a precise, but implementation-independent language. They can be very useful to fundament and verify computational models (or at least part of them) and to compare existing computational models of emotions. Among the best known are formalizations of psychological models by set theory (Broekens *et al.*, 2008; Reisenzein & Junge, 2012) and using agent logics (Hughes & Cresswell, 1996). The latter are the most popular, since, although they are more restrictive in the form of representation, they are usually more useful for a specific application context. One of the agent logics most used is based on BDI terms (belief–desire–intention), which are known as BDI logics. Those are formal logic languages that arise from the combination of several modal logics: a temporal or a dynamic logic used to capture the dynamic nature of agents, and logics for the mental states of belief, desire and intention (Reisenzein *et al.*, 2013).

Specialized affective architectures are those specifically designed for the creation of affective agents. They are composed of several modules (perception, memory, appraisal and actuation) and they are suited to model different affective theories. In this regard, the affective architectures FLAME (El-Nasr *et al.*, 2000), FAtiMA (Dias *et al.*, 2014), Cathexis (Velásquez & Maes, 1997), ALMA (Gebhard, 2005) and MAMID (Hudlicka, 1998, 2004) stand out.

Extended cognitive architecture systems make use of existing general cognitive architectures, such as Soar and ACT-R, or existing agent's architectures such as BDI to model emotions. Their use has the practical advantage of reliving the modeler of the need to solve numerous implementation problems and focus on modeling emotions. Besides, they may help to bridge the gap between psychological emotion theories and computational emotion modeling, as well as they may increase the comparability of different computational emotion models. EMA (Marsella & Gratch, 2009), Sigma (Rosenbloom, 2013) and the proposal of Pérez *et al.* (2016) are based on Soar. Other works such as Fum and Stocco (2004) and Cochran (2006) are extensions of the ACT-R architecture. The emotional extensions of the BDI architecture are often called Emotional BDIs or EBDIs and are in the scope of this paper. Besides their psychological foundations, the existence of logical models (BDI logics) that allow to define, and reason about BDI agents in a precise manner make them very interesting for the creation of intelligent affective agents. The existence of several software systems for programming BDI agents that are close to these logical specifications have also greatly contribute to their spreading.

The aim of this paper is to review and structure the literature on BDI emotional agents, and, based on this, to identify common findings, good practices and open research questions. This comprises both those works based on BDI formalizations as well as those that are extensions of the BDI agent's architecture. As is common when emotional issues are tackled, there is a great diversity of approaches to model affective states and capacities. In the next section, the affective theories most used in the modeling of emotional intelligent agents are reviewed.

3 Affect, empathy and emotional regulation

In the literature, the term affect is frequently used as a general category that includes short-term emotional states such as emotions, long-lasting mood and more permanent states such as personality. Modeling affect in intelligent agents has not only focused on the generation of these affective states but also on their influence on the cognitive process. Other aspects that are often included in the development of intelligent agents are the ability to empathize with others and to regulate emotions and behavior. Empathy is the ability to feel, imagine and support the experiences and emotions of another person (Gibson, 2006), develops gradually throughout life and depends on the degree of contact with the person to empathize. In the field of intelligent agents, empathy has been used as a powerful tool to improve their cooperative and communicative social skills with human users. Regarding regulation, there are two fundamental approaches: emotional regulation and coping. Emotional regulation refers to the way individuals modify the emotions they experience and how they express them. Coping is oriented to affront stressing situations.

In this section, the main affective theories used to develop affective intelligent agents will be reviewed.

3.1 Emotion

In the last decades, diverse theories have demonstrated the influence of emotion in the cognitive process (Damasio, 1994). The definition of emotions is a controversial issue since there are more than 92 definitions in the literature (Kleinginna & Kleinginna, 1981). In general, it is said that emotions can be defined as the 'evaluative judgments of the environment, the self and other social agents, in light of the agent's goals and beliefs' (Hudlicka, 2008). Emotions are short-lived affective states that arise suddenly in the context of a specific stimulus, having significant psycho-corporeal repercussions and producing physiological changes. Three fundamental characteristics of emotions are their high intensity, their short duration and that they are tied to a particular event. There are two principal aspects to take into account in emotion modeling: emotional representation and emotional elicitation.

Regarding emotional representation, some dimensional theories have been used in intelligent agents modeling such as Russell (1980) and Mehrabian (1996b). In these models, affective states are not independent from one another, rather, they are related to one another in a systematic manner (Gunes *et al.*, 2011). In the case of Russell (1980) is a bidimensional model that represents emotional states according to two universal values: valence and arousal. The proposal of Mehrabian (1996b) presents the PAD theory that is a tridimensional model, used not only for representing emotional states, but also for representing mood (see Section 3.2).

Emotional elicitation is the most studied aspect in emotional agent modeling; therefore, several theories of emotion have been used (Frijda, 1987; Zajonc *et al.*, 1989; Ortony *et al.*, 1990; Roseman *et al.*, 1990; Smith & Lazarus, 1991; Damasio, 1994; LeDoux, 1996; Sherer, 1999; Plutchik, 2001). In general, according to the emotion elicitation process, they can be grouped into cognitive and non-cognitive theories (Reisenzein *et al.*, 2013). In general, non-cognitive theories of emotions follow two main perspectives regarding the elicitation of emotions. There are theories that claim that certain kinds of affect (i.e., sensory-pleasure for smelling a rose) are non-cognitively generated, because they presuppose only non-propositional or even non-conceptual representations (Reisenzein, 2006). Other theories claim that even prototypical emotions (i.e., fear, anger, joy) can be non-cognitively caused, given that they can be elicited by the visual or acoustic appearance of danger (i.e., visualize a fierce animal), without mediating any thoughts (Reisenzein *et al.*, 2013). These theories have been less used in intelligent agent modeling, being the cognitive ones the most used.

In cognitive theories, the emotion elicitation requires certain 'higher-order' mental representations (beliefs and desires) (Green & Green, 1992). In the case of non-cognitive theories, emotions in a broad sense (i.e., sensory affects-hedonic pleasure, fear) are always elicited by stimuli via a more direct route that circumvents a higher cognitive processes. Cognitive theories are the most used and in most of them a cognitive evaluation process (appraisal) of the stimulus precedes emotion (Frijda, 1987; Ortony *et al.*, 1990; Roseman *et al.*, 1990; Smith & Lazarus, 1991; Sherer, 1999). Cognitive appraisal theories attempt to explain why a certain event results in one emotional response rather than another and why a certain emotion can be elicited by different events. The key concept is that the subjective cognitive evaluation of events in relation to the agent's goals is responsible for emotion (Roseman & Smith, 2001). More generically, events have to be evaluated as having personal meaning or relevance (Van Reekum, 2000). This evaluation is the so-called appraisal. The most popular of the cognitive theories is the OCC theory (Ortony *et al.*, 1990). This is due to its tree structure that allows implementing it in a simple way, and also because it works with concepts (beliefs, desires and standards) well-studied in agents theories. It is important to emphasize that the OCC theory does not specify the influence of the emotions in the cognitive process, but only their elicitation. The theory bases its foundation on the way individuals perceive the world and it defines three aspects in which the individual can focus its perception: events, agents and objects. If the focus is put on the consequences of events, fortune-of-others, prospect-based and well-being emotions may be elicited. If the focus is put on the actions of the agents, attribution emotions may be elicited. If the focus is put on the properties of objects, attraction emotions are obtained. This way, OCC defines 22 emotions describing their elicitation as a result of the cognitive appraisal process of the stimulus, guided by some evaluation variables (appraisal variables). The theory also defines the global and local variables that influence the intensity of the emotions. As said before, the success of this theory lies in its easy transfer to a computational model since it explicitly defines the elicitation of the emotions and the appraisal variables for its evaluation, as well as the detailed qualitative and quantitative description of both processes. Figure 1 shows the overall structure of this theory.

An interesting aspect of the modeling of emotion is that they have been usually classified into three main types: primary, secondary and tertiary emotions (Damasio, 1994; Sloman, 1998). Primary emotions (being startled, terrified, sexually stimulated) are innate responses to an external event. Secondary emotions (anxious, apprehensive, relieved, pleasantly surprised) generation involves primary emotions and a more complex cognitive appraisal. Tertiary emotions are disturbing states (feeling humiliated, infatuated, guilty) that reduce self-control. Primary emotions generation has usually been modeled on intelligent agents using a cognitive reactive layer, whereas secondary emotions require a deliberative cognitive layer in addition to the reactive machinery of the primary emotions. The case of tertiary emotions is more complex and has not yet been successfully implemented since it depends on the existence of a mechanism that controls the partial loss of control of attention in the intelligent agent, requiring a sophisticated information processing architecture (Sloman, 2000).

Another interesting question is the ability to measure the intensity of which an emotion is experienced since it is a quantitative measure that defines a line between affective reactions and emotions. According to Ortony *et al.* (1990), an emotion only occurs when the intensity achieves a specific threshold; this means that an intelligent agent not always will experiment an emotion. The intensity also determines how strong the response of an individual is (Scherer, 2000). An affective state that depends on the intensity of the emotions is mood, which will be discussed in the next section.

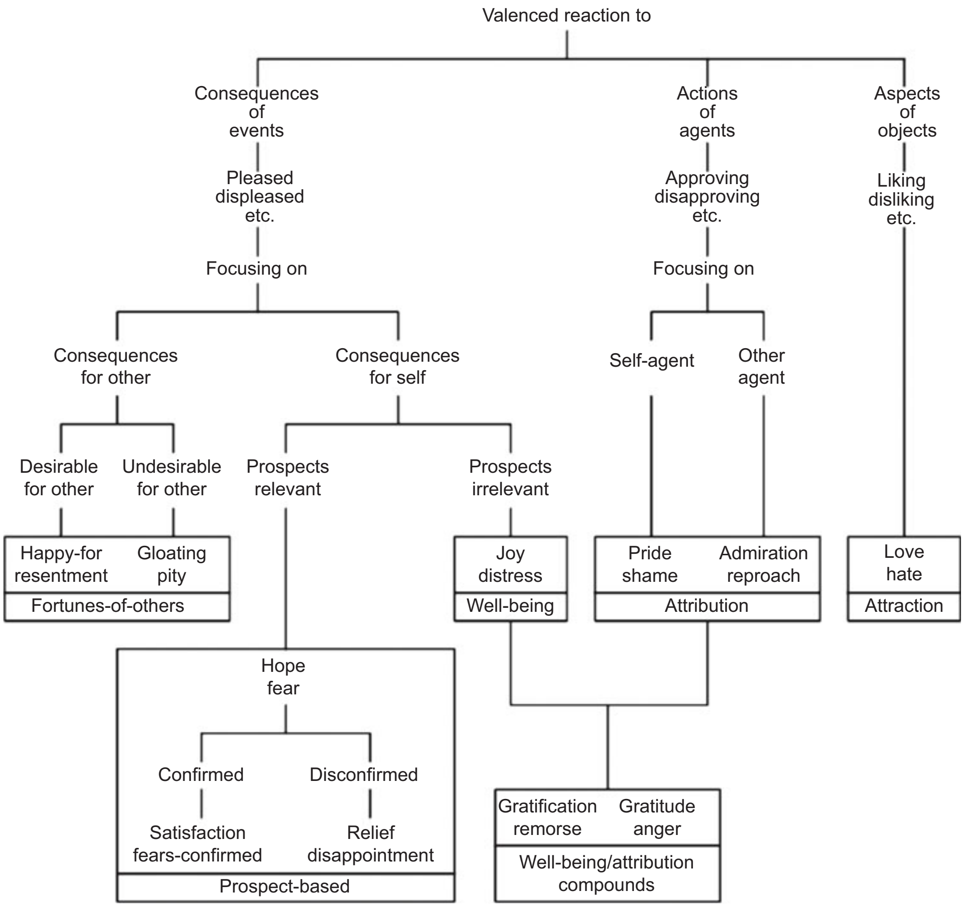

Figure 1 Global structure of the emotions of the OCC theory. Taken from the book 'The Cognitive Structure of Emotions' (Ortony *et al.*, 1990)

3.2 Mood

Mood is recognized as the global affective state of an individual (Desmet, 2015). This affective state tends to vary in time due to emotions, returning afterward to its initial state. Compared with emotions, mood lasts longer, has lower intensity and often appears without being associated with a specific event or object (Larsen, 2000). Gratch and Marsella (2004) refer the importance of modeling moods as 'they have been shown to impact a range of cognitive, perceptual and behavioral processes, such as memory recall (mood-congruent recall), learning, psychological disorders (depression) and decision-making' (Gratch & Marsella, 2004). Several mood theories can be mentioned such as: Larsen (2000), Mehrabian (1996b), Watson and Tellegen (1985) and the PAD theory (Mehrabian, 1996b).

Most of the theories model mood in a bidimensional way, by means of two edges universally recognized as fundamental to characterize mood values: valence and arousal. In this regard, the PANAS theory (Watson & Tellegen, 1985) and the Circumplex theory of affect (Russell, 1980) stand out. In the case of Russell theory, (Russell, 1980) postulates that the underlying structure of an affective experience can be characterized as an ordering of affective states on the circumference of a circle (Remington *et al.*, 2000). Sixteen mood states are defined in terms of valence and arousal. The valence variable is related to negative or positive affect and the arousal value measures the excitement level. The PANAS theory proposed by Watson and Tellegen (1985) does not specify particular mood states but mood categories. Mood states are also defined in two dimensions: valence (pleasant–unpleasant) and arousal (high energy–low energy). Four categories of mood arise in this model: energized–unpleasant, energized–pleasant, calm–unpleasant and calm–pleasant. In the case of the PAD model, mood is modeled in a tridimensional way. Pleasure–Displeasure that is an evaluation of predominance of positive versus negative affective states; Arousal–Non-arousal is the level of physical activity and mental alertness; and Dominance–Submissiveness that is a measure of feelings of control versus lack of control. The PAD theory is the most widely used in intelligent agents.

Mood is generally considered influenced by personality, commented in the next section.

3.3 Personality

Personality refers to characteristic thought patterns that persist through time and situations and that distinguish one person from another. It also influences the emotional response of each individual as well as its mood. Personality and emotions are different in terms of focus and duration. Emotions are always tied to a particular event or object and are short term. Personality remains constant in time and is not tied to a particular event. It can be said that an emotion is a specific and short-lived modification of personality.

In the field of intelligent agents, the most widely used personality theory is the Big Five theory (McCrae & John, 1992), also known as Five Factor Model (FFM). The model establishes five factors that are universally considered as the basic dimensions of personality: openness, conscientiousness, extraversion, agreeableness and neuroticism, frequently known out under the acronym OCEAN. The model has been implemented in several intelligent agents' proposals (André *et al.*, 2000; Yu & Choi, 2005; Durupinar *et al.*, 2008). One popular example is ALMA (Gebhard, 2005) where the personality is used to calculate the initial mood of the agent. To do so, the work of Mehrabian (1996a) that proposes a mapping between PAD and Big Five is used.

There are other personality theories such as The four temperaments theory (Merenda, 1987), the P-E-N theory (Eysenck, 1967) and the Interpersonal Circumplex or IPC model (Wiggins, 2003), but they have been scarcely used in the modeling of intelligent agents. The Circumplex model defines a broad set of interpersonal relationships that have a direct relation with behavior in the affective and cognitive area. It is a bidimensional model that defines two orthogonal axes: a vertical axis (of status, dominance, power or control) and a horizontal axis (of solidarity, friendliness, warmth or love). The four temperaments model is one of the oldest personality models in the world and it is also known as the Hippocrates–Galen temperaments theory (Merenda, 1987). The model defines four temperaments: sanguine, choleric, melancholic and phlegmatic. Another interesting personality theory is the P-E-N (Psychoticism–Extraversion–Neuroticism) of Eysenck (1967). The psychoticism trait is related to normality which means that individuals with higher *p* scores are prone to engage in irresponsible behavior (i.e., aggression, egocentrism and impulsiveness) and to contravene social norms, motivated by a need for immediate gratification, in spite of its consequences. Extraversion trait is modeled according to the two extremes of spectrum, that is, introversion and extroversion, according to the related concepts of inhibition and excitation. Neuroticism is explained with regard to the psychological concept of emotionality where high scores on N factor reveals a low threshold of emotionality, while lower scores indicated a high threshold of emotionality (Eysenck, 1991). This theory is considered together with the Big Five theory, one of the 'giants' in personality modeling (Scholte & De Bruyn, 2004), but in spite of this, it has been scarcely used in intelligent agents.

3.4 Empathy

The most common definition of empathy is 'put yourself in someone else's shoes to understand his emotions' (Pacherie, 2004). An important feature of empathy is that it has two dimensions: cognitive and affective. The cognitive dimension is about understanding the situation of the other and the affective one is about feeling an emotion produced by the other's situation. Both may or may not appear together, it means that an empathic cognitive emotion can be felt without feeling an affective one. The modeling of empathy in intelligent agents has been considered from two perspectives: the empathy evoked by the intelligent agent (into the user) or the empathy that the intelligent agent 'feels'. There are proposals in both directions: those that seeks to evoke empathy in users (Paiva *et al.*, 2004) and those that focused on the empathy that agent feels (Reilly, 1996; McQuiggan & Lester, 2006; Ochs *et al.*, 2008). The latter ones are those of interest here. Usually, these works are not based on specific models but on empirical works or theoretical approaches such as the mentioned OCC theory (Ortony *et al.*, 1990). In fact, OCC describes the elicitation of two empathic emotions when a goal of another intelligent agent or user is achieved: happy-for and sorry-for. To consider them, the intelligent agent needs to have a representation of other agent's (or user) goals.

3.5 Emotional regulation and coping

Emotional regulation is defined as 'the processes by which individuals influence which emotions they have, when they have them, and how they experience and express these emotions' (Gross, 1998). Other authors define it as 'the extrinsic and intrinsic processes responsible for monitoring, evaluating and modifying emotional reactions, especially their intensive and temporal features, to accomplish one's goals' (Thompson, 1994).

Another concept regarding regulation is coping; its more accepted definition is given by Lazarus and Folkman (1984) who express that coping 'consists of cognitive and behavioral efforts to manage specific external or internal demands (and conflicts between them) that are appraised as taxing or exceeding the resources of the person'. It should be noted that emotional regulation focuses on the increase or decrease of positive and negative emotions (Gross, 1998), whereas coping is more about diminishing negative emotions in stressful situations (see Lazarus & Folkman, 1984; Scherer, 1984). The most used models in the field of agent modeling are the one of Smith and Lazarus (1991), oriented to coping and the Gross theory (Gross, 1998) oriented to emotional regulation.

In Smith and Lazarus (1991), coping is related to the way the individual interprets its relation with the environment. The theory proposes two fundamental coping strategies: problem-focused, that consists of altering the adverse situation, and emotion-focused, that consists of coping distressing emotions that arise in situations where the circumstances not going to change. An example of the use of this model in intelligent agents is EMA (Marsella & Gratch, 2009), a proposal to create agents that are able to cope with negative affect.

In Gross (1998), emotions can be regulated in five different phases during the emotion generation process: Situation Selection (avoidance of people, places or objects), Situation Modification (active efforts to modify the situation in order to alter its emotional impact), Attentional Deployment (selection of where to put the focus on), Cognitive Change (modification of the cognitive evaluation of the situation) and Response Modulation (expression of the emotion elicited). Each phase takes part in the process according to the emotional dynamics and it is able to regulate positive as well as negative affect. CoMERG (Bosse *et al.*, 2017b) and S-model (Soleimani & Kobti, 2012) are two computational models of emotion regulation that follow Gross's paradigm.

Both coping and emotional regulation are oriented to the regulation of emotions but there are also theories to regulate more medium-term affective states like mood, such as the theory of Larsen (2000). In fact, the affective regulation term has been proposed (Gross, 1998) as a global category that includes emotional regulation, coping and mood regulation. Moreover, there are works that address the subject of empathic emotion regulation (EER) and define it as 'an interpersonal regulation system in which an empathic response to another person's emotional state aims to regulate both emotion within the provider and across interaction partners' (Hamburg *et al.*, 2014). Nevertheless, this kind of issues has not been modeled yet in intelligent agents.

3.6 Theoretical bases of the affective influence on cognitive process

The influence of emotions on the cognitive process is a field of interest in many research areas (philosophy, psychology, neuroscience) (Blanchette & Richards, 2010; Angie *et al.*, 2011). Studies have shown that emotions influence the human decision-making process (Damasio, 1994) and may influence all the phases of the cognitive process: perception (Zadra & Clore, 2011), beliefs (Frijda *et al.*, 2000a; Glore & Gasper, 2000) and behavior (Dolan, 2002).

Emotions, mood and personality are the affective aspects whose influence on the cognitive processes has been studied. As commented before, emotions and mood are two strongly related concepts. This is why in many works about emotional influence, they are treated interchangeably (Zadra & Clore, 2011).

Regarding the influence of emotions on the cognitive processes, there is no general consensus among researchers. There are studies that show that emotions influence the mediating information processing (attention, perception, memory), situation assessment, decision-making, goal management, planning and learning (Frijda *et al.*, 2000a; Dolan, 2002; Angie *et al.*, 2011). For example, positive emotions induce global focus, whereas negative emotions induce a moral local focus and analytical thinking. Also, anxiety states reduce attentional and working memory capacities and biases attention toward the detection of threatening stimuli (Jeon, 2017).

Regarding the influence of mood, it has been studied its impact on a range of cognitive, perceptual and behavioral processes, such as memory recall (mood-congruent recall), learning, psychological disorders (depression) and decision-making (Gratch & Marsella, 2004). Particularly, its influence on judgment is mediated by the individual's subjective (positive or negative) experience about the object of judgment. This means that, when evaluating objects or belief statements, individuals act accordingly to how they feel about them, using that information for the judgment in a mood-congruent way (Schwarz & Clore, 1983).

Regarding personality, studies have been focused mainly on its influence on behavior (McEachan *et al.*, 2010; Lönnqvist *et al.*, 2001). For example, a growing number of researchers have concluded that neurotic (as compared to non-neurotic) individuals have stronger negative affective reactions to undesirable stimuli, whereas extraverts (as compared to introverts) have stronger positive affective reactions to pleasant stimuli (i.e., Larsen & Ketelaar, 1989; Larsen & Ketelaar, 1991).

In the Section 5, the BDI proposals that deal with the affective aspects studied in this section will be reviewed. First, the main BDI concepts are reviewed.

4 The BDI agent's architecture

BDI architecture is born on the basis of the theoretical foundation of Michael Bratman's practical reasoning (Bratman, 1987). Until then, there were models that only considered beliefs and desires; BDI introduces future-direct intentions providing the deliberative character to the model. The intentional feature was introduced by Bratman and is inspired by the 'intentional instance' term of the philosopher Daniel C. Dannett (1971). BDI is known for its simplicity and efficiency in shaping human behavior and is based on three fundamental mental attitudes:

- *Beliefs*: that represent the information about oneself and the environment.
- *Desires*: that represent the motivational state.
- *Intentions*: that represent the commitments to achieve desires and provide the deliberative character to the model.

In Figure 2 shows the interaction between the three main components of the BDI:

- When an event happens (either internal or external), it is perceived by the agent through the input sensors (*pers(Env)*).
- With new perceived information, beliefs are updated (*brf(B,p)*).
- Based on the set of updated beliefs, the desires are generated (*options(B,I)*).
- Intentions are obtained to fulfill the objectives outlined (*filter(B,D,I)*).
- Finally, plans of actions are executed (*action(I)*).

BDI offers several features that make it popular:

1. Beliefs provide a powerful tool to remind past events.
2. Plans are not fully defined and may consist of a set of sub-objectives, which provide adaptability to the framework.

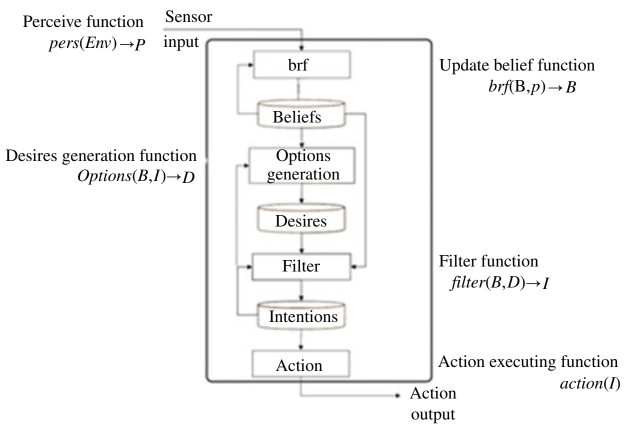

Figure 2 Generic BDI schema adapted from Weiss (1999)

3. The reconsideration of the plans that are being executed allows dealing with a dynamic environment because it allows to meet more urgent needs or to discard intentions that have no reason to be.
4. It is based on goals and not on tasks. This allows knowing at every moment not only *what* but *why* the agent is doing something. This way it is possible to explain its behavior at every moment, giving the system the possibility to recover from failures and seize opportunities that dynamically arise.
5. The programming is at a high level and allows the development of more complex systems in a simple way.

Anand S. Rao and Michael P. Georgeff are among the first developers to put the BDI architecture into practice implementing an air traffic management application, an air combat simulation system (Rao & Georgeff, 1992, 1995) and an autonomous mobile robot (Georgeff & Lansky, 1987). Since then, BDI agents have been created and used in several application such as the tourist guide of Corchado *et al.* (2004) or the virtual humans training system of Van Oijen *et al.* (2011).

The growing popularity of the BDI architecture has led to the development of several implementation tools that support it. The first generation of languages that emerged was procedural reasoning system (PRS) (Ingrand *et al.*, 1992), followed by its successor dMARS (D'Inverno *et al.*, 2004) continued by its third generation JACK Intelligent Agents (Winikoff, 2005) and AgentSpeak (L) (Rao, 2009) that emerged as a new approach to formalize the operational semantics of PRS and dMARS. The AgentSpeak (L) language is one of the most influential languages in the implementation of BDI agents and has been developed up to its free version implemented in Java, JASON (ref), which, being an open source interpreter, allows it to be completely configurable and easy to implement. Other languages that also support BDI agents are JAM (Huber, 1999), 3APL (Hindriks *et al.*, 1999), SPARK (Morley & Myers, 2004) and JADEX (Braubach *et al.*, 2003). In fact, JADEX is one of the most attractive languages due to its adaptability to run on several platforms, in addition to being implemented in Java and also being an open source.

In the last years, the growing interest in incorporating affect into intelligent agents has given rise to several emotional BDI proposals that are reviewed in the next section.

5 Review of emotional BDI proposals

In this section, we will present the most relevant proposals related to the modeling of affective aspects in BDI agents. A systematic bibliographic search was done. From the 978 works initially obtained, we restricted them to those published in the last 20 years in English that presented original BDI proposals (no revision or comparisons or new applications), prioritizing those with a medium or high relevance (in terms of citations).

In the first section, we present a general classification of the selected works. Then we review them in terms on how the different affective aspects are considered and how they influence the general cognitive BDI process.

5.1 Classification of the emotional BDI proposals

Following the classification presented in Section 2, the proposals found can be grouped into two main categories: BDI formalizations and extensions of the BDI architecture. The works are presented in Table 2 indicating the affective aspect they consider (emotions, mood, personality, empathy and regulation). The level of application is also indicated: some proposals are no implemented at all, other present just an experimental scenario or application to test its viability. Only a few present a full application with a specific objective.

The first group of proposals is formed by those belonging to the formalization group. It is quite a diverse group. Most of them are based on agent logics languages, in particular, in BDI logics. But others extend existing agents specification frameworks or agent modeling languages as it is commented below.

The proposals of Steunebrink *et al.* (2012) and Meyer (2006) extend the BDI-based KARO agent specification framework (Meyer *et al.*, 1999). The KARO framework is based on a blend of dynamic logic with epistemic logic, enriched with modal operators for beliefs and desires concepts (Meyer & Van Der Hoek, 2004). No application or scenario is presented.

The works of Adam *et al.* (2009), Dastani and Lorini (2012) and Gluz and Jaques (2017) use BDI logic languages: the ABC, Dynamic Logic of Graded Attitudes (DL-GA) and AfPL logics, respectively. The ABC logic is an extended version of one of the first BDI logics proposed by Cohen and Levesque (1990). It formalizes the concepts of action, time, belief and choice in a simple propositional modal logic (Herzig & Longin, 2004). The DL-GA logic is designed to represent with special operators the beliefs, goals and intentions of the agent, where beliefs and goals have degree of plausibility and desirability, respectively. The AfPL logic is a propositional logic with the traditional logic operators that manage the beliefs, goals intentions concepts.

As we will see in the next section, in all there cases, they formalize the OCC theory, combined with Oatley and Jenkins (1996) theories in the case of Adam *et al.* (2009) and the Lazarus (1991) and Frijda (1986) theories in the case of Dastani and Lorini (2012). Agent logic languages facilitate the transformation from specification to implementation. Nevertheless, Gluz and Jaques (2017) is one of them that presents an experimental scenario (a vacuum cleaner robot), but without implementation.

Finally, in this group, the proposals of Bosse and de Lange (2008), Bosse and Zwanenburg, (2009), Bosse and Höhle (2011) can also be found. In this case, they use as formalization tool an agent-based modeling language, called LEADSTO (Bosse *et al.*, 2007a). LEADSTO language allows to model dynamic processes in terms of both qualitative and quantitative concepts, by the means of either logical or mathematical means, or a combination thereof. They formalize three Ekman's emotions, the expectation-based emotions of Castelfranchi and Lorini (2003) and two emotions based on the appraisal theory presented in Gratch and Marsella (2004). They present three experimental applications (simulation environment with two IVA, a dice game with an embodied agent and game based on RobCop, respectively) to test the viability of the proposals.

As we see, these works have focused mainly on emotional formalization and mostly on formalizing the OCC theory. Only three of them have considered other aspects such as mood (Bosse & Höhle, 2011) and regulation (Bosse & Zwanenburg, 2009; Dastani & Lorini, 2012). Details on how they consider these three aspects are presented in the next section. Empathy and personality have never been formalized.

The second big category is formed by those works that are extensions of existing BDI cognitive architectures. They follow the generic BDI schema (see Section 4) extended with affective modules that influence different cognitive phases (as will be explained in Section 5.7). Although OCC is also widely used, many more theories are applied (as it will be commented in the next section). Mood and personality are also considered in about half of them. On the contrary, empathy (just treated in three of them) and regulation (only in two) are scarcely considered.

In these cases, as it was commented in Section 4, the existence of several implementation tools facilitates their development and application. In spite of this, some of them do not present any application or example scenario. Other just present experimental applications or scenarios to test the architecture's viability. This is the case of Jiang *et al.* (2007) that present an application developed in the Tileword platform (Pollack & Ringuette, 1990) to compare the behavior of emotional and non-emotional BDI agents. Puica and Florea (2013) present an experimental scenario of an agent situated in a 2D map with artifacts and traps. Hernández *et al.* (2004) present a more elaborate experimental application based on a humanoid robotic head with vision and audio sensor capabilities to evaluate the reaction of the robot to certain colors and the activation of primary emotions (surprise, fear) and secondary emotions (happiness, sadness and anger) given the image's luminance value. Neto and da Silva (2012) present two games as experimental applications: iterated prisoner's dilemma and a memory card game. They focus on evaluating the influence of the affective state in the activities of perception, memory and decision-making. In the case of Boukricha and Wachsmuth (2011), the experimental application is an embodied agent, called EMMA, integrated as third pattern into Max's virtual environment (Becker-Asano & Wachsmuth, 2008). They focus on the assessing EMMA's capability of empathizing with Max's emotions during the interaction with a human user. Four of the proposals present full applications. In all the cases, an embodied conversational agent is used but within different application scenarios. De Rosis *et al.* (2003) work has been used as the base of an interactive listening agent (Bevacqua *et al.*, 2008) and an Interactive Storyteller (Bevacqua *et al.*, 2010). In Ochs *et al.* (2010), the application is 3D talking head that shows facial expressions according to the intensity of the elicited emotions, in order to empathize with the user. In the case of Boukricha and Wachsmuth (2011) and Becker-Asano and Wachsmuth (2008), they present two different systems, the first one focused on empathic emotion elicitation and the second one on primary and secondary emotion elicitation. The application developed by Becker-Asano & Wachsmuth (2008) is a virtual human named Max that plays the role of museum guide. The application has been implemented in several scenarios (robotics, chat boot agent) and also more recently it has been extended with memory retrieval functions of the ACT-R cognitive architecture (Becker-Asano, 2014).

In Table 1, the reviewed emotional BDIs are presented showing the affective aspects considered and their level of application.

5.2 Emotion modeling

Emotion is the most commonly treated emotional concept in intelligent agents. As it has been analyzed in Section 3.1, there are many emotion elicitation theories that have been used in this context and they can be grouped into cognitive and non-cognitive theories. In cognitive theories, the emotion elicitation requires certain 'higher-order' mental representations whereas in the case of non-cognitive theories, emotions are elicited by stimuli via a more direct route that circumvents higher cognitive processes (Reisenzein *et al.*, 2013; Reisenzein *et al.*, 2013). The emotional BDI proposals selected to be reviewed will be classified into those two groups.

Only three proposals follow non-cognitive theories. In this group, the work of Jiang *et al.* (2007) stands out and is the basis of other proposals. They follow the theory of Damasio (1994) but the concept of emotion is implemented in general way, without specifying the elicitation of any particular emotion. They provide the cognitive affective reasoning needed for eliciting both primary and secondary emotions in the general BDI process. Primary emotions are modeled as the first feeling that an individual experiences when an event occurs and secondary emotions may be caused directly by primary emotions, or come from more complex chains of thinking. Current beliefs and intentions originate primary emotions, and the generation of secondary emotions is influenced by beliefs, intentions and primary emotions. In contrast with the other examples, beliefs can be acquired through perception, contemplation or communication. In the case of Jones *et al.* (2009), they maintain the cognitive and emotional modeling of Jiang *et al.* adding personality and physiology, which is their main contribution. Physiology (stress, hunger/thirst, tiredness, temperature, injury and contamination) is used as a moderator of the cognitive-affective process that directly influences the elicitation of primary emotions. Personality modeling is another important aspect that differences these proposals as it will be explained in Section 5.4.

In the case of Bosse and de Lange (2008), their modeling of emotions is based on Ekman *et al.* (1972) and just three basic human emotions (anger, happiness or fear) are elicited. In contrast to the other proposals in this group, beliefs can only be acquired from observations and only one emotion can be experienced simultaneously. Besides, each emotion has an associated emotional response level (ERL) directly influenced by agent's beliefs. In spite of the simplicity of their emotional modeling, they address

Table 1 A classification of the proposals reviewed in this paper showing the affective aspect that model (mood, personality and regulation) and the application type (Exp_app: Experimental application; Exp_scenario: Experimental scenario; Real_app: Real application)

	Proposals	**Emotions**	**Mood**	**Personality**	**Empathy**	**Regulation**	**Application type**
BDI formalizations	Adam *et al.* (2006)	√	–	–	–	–	–
	Steunebrink *et al.* (2012)	√	–	–	–	–	–
	Gluz and Jaques (2017)	√	–	–	–	–	Exp_scenario
	Meyer (2006)	√	–	–	–	–	–
	Dastani and Lorini (2012)	√	–	–	–	√	–
	Bosse and Zwanenburg (2009)	√	–	–	–	–	Exp_app
	Bosse and de Lange (2008)	√	–	–	–	√	Exp_app
	Bosse and Höhle (2011)	√	√	–	–	–	Exp_app
Extension of BDI architectures	Pereira *et al.* (2005)	√	–	–	–	–	–
	Van Straalen *et al.* (2009)	√	–	–	–		–
	Alfonso *et al.* (2014)	√	√	√	–		–
	Van Dyke Parunak *et al.* (2006)	√	–	–	–	–	–
	Jiang *et al.* (2007)	√	–	–	–	–	Exp_app
	Lejmi-Riahi *et al.* (2014)	√	–	–	–	–	–
	Hernández *et al.* (2004)	√	√	–	–	–	Exp_app
	Jones *et al.* (2009)	√	–	√	–	–	–
	Puica and Florea (2013)	√	–	√	–	–	Exp_scenario
	Neto and da Silva (2012)	√	√	√	–	–	Exp_app
	Boukricha and Wachsmuth (2011)	√	√	√	√	–	Real_app
	Ochs *et al.* (2010)	√	–	–	√	–	Real_app
	De Rosis *et al.* (2003)	√	√	√	–	√	Real_app
	Becker-Asano and Wachsmuth (2010)	√	√	√	√	√	Real_app

Exp_app = Experimental application; Exp_scenario = Experimental scenario; Real_app = Real application an interesting and practically unexplored field in intelligent agents modeling, emotional regulation, as it will be analyzed in Section 5.5.

The proposals based on cognitive theories are the more numerous and have been divided into formalization and extension proposals. In the formalization category, four of them are based on the OCC theory and the others in Lazarus (1991), Oatley and Jenkins (1996) and Castelfranchi and Lorini (2003), respectively.

The proposals that formalize the OCC theory have in common the way emotions are elicited, because they consider that emotions depend on beliefs and—directly or indirectly—on desires. In Adam *et al.* (2009), emotion elicitation involves two types of beliefs (strong beliefs, that are those that are truth without doubt, and weaker beliefs, that are those that are related to the probability of an event to be true) and goals that are related to the desirability/undesirability of the consequences of events. They manage intensity, but they do not deferenciate between emotion elicitation and emotion experience. This means that in the case that the conditions for an emotion are met, the agent will always experience it. In addition, they do not distinguish between emotions of the same OCC branch as anger and irritation. Finally, their formalization is restricted to their own BDI-based logical framework. The proposal of Steunebrink *et al.* (2012) is a first attempt to differentiate between affective reactions and experienced emotions: an emotion triggered is experienced only if it has already been experienced in the past and the overall emotion intensity is positive. To do so, they implement an intensity assignment process but without detailing the complete intensity calculation. In contrast to Adam *et al.* (2009), their formalization of the OCC theory it is not tied to a particular framework, because they provide a formalization into dynamic doxastic logic an later they translate it into a logical implementation. In contrast with the precedent works, Gluz and Jaques (2017) present a probabilistic formalization of the OCC emotion elicitation process. They also distinguish between emotion elicitation and emotion experience managing intensity, but providing all the detailed calculations of the variables affecting intensity (potential and threshold). As the other works, they formalized the process of appraisal that depends on agents' beliefs, goals and desirability but adding a new desirability operator. In the proposal of Dastani and Lorini (2012), emotions are defined in terms of graded beliefs, graded goals, and intentions and intensity is considered. Their implementation of coping strategies represents a step forward in including affective aspects into formalizations and it will be discussed in Section 5.4. All the precedent proposals formalize OCC emotions but they differ in the number of emotions they consider: Steunebrink *et al.* (2012)-22 OCC emotions, Adam *et al.* (2009)-20 emotions, Gluz and Jaques (2017)-8 emotions (event-based emotions with consequences for self-based) and Dastani and Lorini (2012)-4 of the prospect-based emotions (see specific emotions in Table 2).

In the formalization group, there are also proposals based on different theories of emotion, not OCC. This allows them to model other types of emotions not usually considered. The proposal of Meyer (2006) is based on the emotional theory of Oatley and Jenkins (1996) and models the elicitation of free-floating emotions, whose elicitation is not always directly related to a particular object or situation. They formalize four types of them: happiness, sadness, anger and fear. The elicitation of emotions is shaped by the success or failure of a plan and by the achievement or abandoning of the pursued goal. They do not manage the intensity of emotion. Bosse and Zwanenburg (2009) proposal is based on the emotional theory of Castelfranchi and Lorini (2003) and considers expectation-based emotions (before-emotions and after-emotions). Modeling this kind of emotions allows the agent not only to 'feel' an emotion when an event happens, but also by anticipating future events (before-emotions) and evaluating past events (after-emotions). Expectation-based emotions are generated by the combination of beliefs and desires in a particular state, considering also the intensity related to the importance of the desire. The proposal of Bosse and Höhle (2011) is based on Gratch and Marsella (2004) work. Emotions are generated as a result of the evaluation of events through desires and are influenced by beliefs. Two types of emotion are modeled, happiness and sad, but the model can also be used to represent emotions like anger and fear. They do not differentiate between emotion elicitation and emotion experience, but they work with the concept of intensity proportional to the strength of the desire. Their proposal differentiates from all the other because they include mood modeling as it will be discussed in Section 5.3.

Table 2 The proposals reviewed showing the theory in which the emotional modeling is based on and the emotions they model

	Proposal	Category	Emotional theory	Emotions modeled
Non-cognitive	Bosse and de Lange (2008)	F	Ekman *et al.* (1972)	Anger, happiness and fear
	Jiang *et al.* (2007)	Ex	Damasio (1994)	Primary and secondary emotions unspecified
	Jones *et al.* (2009)	Ex	Damasio (1994)	Fear/hope, anger/gratefulness, shame/proud and reproach/trust
Cognitive	Meyer (2006)	F	Oatley and Jenkins (1996)	Free-floating emotions: happiness, sadness, anger and fear
	Bosse and Zwanenburg (2009)	F	Castelfranchi and Lorini (2003)	Expectation-based emotions: before-event (hope and fear) and after-event (surprise, (dis)satisfaction, relief and disappointment)
	Adam *et al.* (2009)	F	OCC theory (Ortony *et al.*, 1990)	Event-based and agent-based emotions
	Bosse and Höhle (2011)	F	Based on Gratch and Marsella (2004)	Happiness and sadness
	Dastani and Lorini (2012)	F	OCC theory (Ortony *et al.*, 1990)	Prospect-based emotions: hope, fear, joy, sadness and distress
	Gluz and Jaques (2017)	F	OCC theory (Ortony *et al.*, 1990)	Event-based emotions with consequences for self-based
	De Rosis *et al.* (2003)	Ex	OCC theory (Ortony *et al.*, 1990)	Event-based emotions
	Van Dyke Parunak *et al.* (2006)	Ex	OCC theory (Ortony *et al.*, 1990)	Event-based (fear) and agent-based (anger) emotions
	Van Straalen *et al.* (2009)	Ex	OCC theory (Ortony *et al.*, 1990)	Unspecified
	Ochs *et al.* (2010)	Ex	Own model of emotions, based on theoretical approach	Satisfaction, frustration, irritation, sadness and anger as empathic emotions

Table 2 Continued.

	Proposal	Category	Emotional theory	Emotions modeled
	Boukricha and Wachsmuth (2011)	Ex	OCC theory (Ortony *et al.*, 1990)	Primary emotions (happy, surprised, angry, annoyed, bored) and empathic emotions related
	Steunebrink *et al.* (2012)	Ex	OCC theory (Ortony *et al.*, 1990)	Event-based, agent-based and object-based emotions
	Neto and da Silva (2012)	Ex	OCC theory (Ortony *et al.*, 1990)	Event-based, agent-based and object-based emotions
	Lejmi-Riahi *et al.* (2014)	Ex	Loewenstein and Lerner (2003)	Immediate emotions and expected emotions unspecified
	Pereira *et al.* (2005)	Ex	Open to any theory	Unspecified
Cognitive and non-cognitive	Hernández *et al.* (2004)	Ex	Damasio (1994) and Sloman (2002)	Primary (fear and surprise) and secondary (happiness, sadness and anger) emotions based on Damasio and tertiary (shame or pride) emotions based on Sloman
	Becker-Asano and Wachsmuth (2010)	Ex	Damasio (1994), OCC theory (Ortony *et al.*, 1990), Ekman *et al.* (1972)	Primary (bored, annoyed and depressed Ekman's basic emotions) and secondary (OCC prospect-based emotions)
	Puica and Florea (2013)	Ex	Damasio (1994), OCC theory (Ortony *et al.*, 1990), Russell (1980)	Primary emotions and secondary emotions based on Damasio unspecified
	Alfonso *et al.* (2014)	Ex	Damasio (1994), OCC theory (Ortony *et al.*, 1990)	Primary and secondary emotions distress, excitement, contentment, depression, misery, pleasure, arousal and sleepiness

F = Formalization; Ex = BDI extension

As most of the proposals in the formalization group are based on BDI logics, they may be not expressive enough to accommodate all assumptions of psychological emotion theories. For example, most of them are purely qualitative; however, emotions differ not only in quality but also in intensity, which is important to explain certain effects of emotions. In this regard, proposals like the one of Dastani and Lorini (2012) that combines a dynamic knowledge logic with a logic for graded beliefs and desires can provide a more wide range of constructs to enrich the modeling of emotions. Most of them miss also another important issue: the emotional influence on cognitive reasoning, that is, how the emotions can help agents to choose the most rational action to be performed, and how the emotions can improve the way that agents reason or decide or act.

Regarding the proposals that are BDI extensions, they are mainly (all except two of the selected proposals) based on the OCC theory, either alone or combined with other theories. Proposals based only on the OCC theory manage the process emotional elicitation in a similar way: an event is appraised in terms of beliefs, desires and intentions and a certain score regarding its valence (negative/positive, displeasure/pleasure) and arousal (intensity of the elicited emotion) is returned. Their main differences consist of emotions they model and the consideration of additional concepts that directly affect the cognitive and affective process. The proposal of Neto and da Silva (2012) is the one that models the 22 OCC emotions. The other proposals focus on modeling a specific branch of the OCC emotions, like the case of De Rosis *et al.* (2003) that model the event-driven category, Van Dyke Parunak *et al.* (2006) that model only two emotions. Other works like Ochs *et al.* (2010) and Boukricha and Wachsmuth (2011) are based on the OCC theory but their main objective is empathy modeling, and therefore will be discussed in Section 5.5.

Among the BDI extensions that are based only on the OCC theory, the proposal of De Rosis *et al.* (2003) is probably the most relevant since it is one of the first emotional BDI proposals and it is one of the most cited. They present an interesting proposal that uses a belief network for representing the agent's mental state. They model the elicitation of a subset of the emotions: the event-driven categories of fortune-of-others, well-being and prospect-based emotions. The elicitation of emotions is determined by the variation in the agent's beliefs about whether a goal can be threatened or archived, depending on the importance (weight) assigned to the goal. They manage the intensity of an elicited emotion depending on the weight of the goal. This idea of assigning weights to desires is similar to the ideas of probability of Gluz and Jaques (2017) and graded goals of Dastani and Lorini (2012). The main problem is the complexity of the proposed beliefs network, since the number of nodes increases considerably with the number of emotions modeled. Nevertheless, they make up for this problem, reducing the time required to propagate evidence and update the emotion intensities, thanks to the effectiveness of the employed algorithm. They also take into account mood, personality and regulation aspects discussed in Sections 5.3, 5.4 and 5.6, respectively.

Some other OCC-based proposals stand out because they consider new concepts, providing the agent with a wide range of information and resources that improve its performance in a dynamic environment. Van Dyke Parunak *et al.* (2006) include the concepts of dispositions, triggers and tendencies. They extend the appraisal reasoning processes with an analysis process. The appraisal process assesses the beliefs in the context of agent disposition and returns the elicited emotion. Beliefs are mapped to digital pheromones, which inform the presence of other agents (or objects) in the environment. These pheromones act as triggers for emotions. The analysis process determines the intentions to be committed by the agent, drawing on the pheromone vector. As it can be seen, their emotional modeling is tied to a particular situation (domain-dependent) and only two emotions are modeled.

Neto and da Silva (2012) manage a memory component for past events and a memory filter that retrieves those beliefs that are aligned with the agent's current emotional state. They also consider that the emotional state of the agent is not only determined by the elicited emotions caused by an event, but also by the agent's perception. To do so, every possible perception has associated a vector of 22 emotions, and the sum of its intensities has to be equal or gather than a minimum value so they can be take it into account in the emotional state of the agent. Another interesting aspect is that the affective output of the emotional component is the agent‘s mood and not the elicited emotions.

As it was mentioned above, only two proposals are not based on OCC theory. Lejmi-Riahi *et al.* (2014) emotional BDI is based on the emotional theory of Loewenstein and Lerner (2003) and is oriented to deal with expected and immediate emotions. Expected emotions are predictions about the emotion that will be experienced if the decision is taken, and immediate emotions refer to the emotions felt at the time of the decision-making. In contrast with the proposals that model OCC, they propose three appraisal process: incidental stimulus appraisal (for immediate emotions elicitation), and emotional option and cognitive option appraisal (for expected emotions elicitation). Immediate emotions generation is influenced by beliefs and intentions and expected emotions generation is influenced by immediate emotions, desires and intentions. The proposal of Pereira *et al.* (2005), unlike all the previous ones, is opened to the implementation of any emotional theory; they do not detail how emotions are modeled, but just the interaction between the emotional component and the BDI process. Their main contribution is the introduction of two interesting concepts: capabilities and resources. The idea of introducing these concepts comes from identifying some disadvantages of the BDI model such as the lack of information about resource bounds, of a dynamic deliberation process and of the consideration of other mental states different than beliefs, desires and intentions. They define capabilities as a library of available abstract plans that the agent owns and resources as the way of converting these abstract plans into action plans. Resources can be either physical (disk space, available memory, etc.), virtual (energy sources on a virtual world, other agents, etc.) or both. The emotional component manages both the resources and the capabilities, and the availability or unavailability of capabilities depends on the availability of the required resources.

The proposals that combine different cognitive and non-cognitive theories of emotion are oriented to model the elicitation and representation of a high diversity of emotions (primary, secondary and tertiary) and to consider additional affective concepts such as mood, personality and regulation. Three of the proposals are based on the Damasio and OCC theories and they model primary and secondary emotions. Of them, the proposal of Becker-Asano and Wachsmuth (2008) stands out for being one of the most cited and has also been used in different applications domains. Their model combines the OCC and the Damasio theory for emotion elicitation and the PAD for emotional representation. Primary emotions are modeled as inborn affective states that are triggered by reflexes in case of harmful stimuli, determining the reactive behavioral responses of the agent. Secondary emotions are modeled as a more complex affective state and their elicitation depends much more on the situational and social context than that of primary emotions. Puica and Florea (2013) focus on simulating human mechanisms of internal resource usage and on agent personality following the ideas of Pereira *et al.* and of Jones *et al.*, respectively. They combine the theory of Damasio for representing primary and secondary emotions, the OCC theory for emotion elicitation and the Russell's theory for emotion representation. Alfonso *et al.* (2014) also follow the theories of emotion of Ortony *et al.* (1990) and Damasio (2005). They include the modeling of the surprise emotion, not considered in OCC, which can be either a primary or secondary emotion. Hernández *et al.* (2004) combines Damasio (1994) with Sloman (2002) theory (not OCC). They propose an emotional component composed of a three-level belief system that generates three emotional levels: primary, secondary and tertiary. Although tertiary emotions are not implemented, this is the first attempt for considering them. They use also model emotional representation in a dimensional space (valence; arousal) based on the ideas of (Breazeal, 2002).

As a conclusion, most of the proposals are based on just one theory (being OCC very prominent) but the idea of combining several theories is interesting and should be further explored. It would allow widening the type of emotions considered to some of them that are practically unexplored, such as Sloman's tertiary emotions. Another interesting aspect is the combination of elicitation and representation emotion theories. In this sense, the OCC theory combined with the PAD model is often used, but other interesting theories such as the Russell and PANAS theories for emotional representation may also be used. Moreover, the consideration of new concepts (resources, capabilities, triggers, tendencies, information not only acquired by perception) provides a wide range of information and resources to the agent and may improve his performance in dynamic environments.

Table 3 Proposals reviewed that model mood showing the affective theory they follow and the mood states they model

Proposal	Affective theory	Moods modeled
De Rosis *et al.* (2003)	–	Not independent state
Hernández *et al.* (2004)	–	Not independent state
Bosse and Höhle (2011)	–	Positive and negative
Becker-Asano and Wachsmuth (2010)	PAD	Positive and negative
Boukricha and Wachsmuth (2011)	De Vignemont & Singer, 2006	Positive and negative
Neto and da Silva (2012)	PAD	Exuberant, dependent, relaxed, sweet, bored, arrogant, anxious and hostile
Alfonso *et al.* (2014)	PAD	Exuberant, dependent, relaxed, sweet, bored, arrogant, anxious and hostile

In Table 2, the revised proposals are presented showing the emotional theory they are based on and the specific emotions modeled. The way these works consider the different affective aspects will be commented in the next section.

5.3 *Mood modeling*

Mood is generally recognized as the global affective state of an individual (Section 3.2). In spite of this, just 7 of the 22 proposals reviewed model mood, being one (Bosse & Höhle, 2011) a formalization and all other extended BDI architectures. In Table 3, those proposals are shown.

As it can be seen, the proposals that deal with mood in the simplest way (De Rosis *et al.*, 2003; Hernández *et al.*, 2004) do not follow any mood theory and do not model mood as an independent state but rather as the decay of the emotions in time. Bosse and Höhle (2011) treat mood as an independent state that depends on the previous mood in combination with the weighted sum of the different emotional intensities, but without applying any explicit affective theory.

In the other four proposals, mood is modeled as an independent state applying an affective theory, being the dimensional one, and more specifically, the PAD theory the most extended. In some cases, mood is modeled directly derived from PAD space (Neto & da Silva, 2012; Alfonso *et al.*, 2014) whereas in other it is modeled as the agent's overall positive or negative feeling (Becker-Asano & Wachsmuth, 2010; Boukricha & Wachsmuth, 2011).

In all proposals, mood persistence in time is modeled as a decay factor that is applied in the absence of another input stimulus. In some cases, the decay factor is determined directly by the agent personality as well as mood's initial or default value (De Rosis *et al.*, 2003; Alfonso *et al.*, 2014). In other cases, the decay factor is determined by a linear function (Becker-Asano & Wachsmuth, 2010), as a change rate (Bosse & Höhle, 2011), or as an amount by which the current emotional state is divided (Hernández *et al.*, 2004).

All proposals also agree in the interconnectivity between mood and emotions which results in 'emotion dynamics': emotions influence mood updating and mood influences emotion elicitation. This is empirically supported by Neumann *et al.* (2001), who found that individuals in a positive mood are less likely to experience negative emotions and vice versa. The influence of mood in the emotional elicitation process is generally named mood congruency. This bidirectional relationship has been particularly modeled in those proposals that model primary and secondary emotions. In most of the cases, mood influences the elicitation of primary emotions (Alfonso *et al.*, 2014) and in some cases, it is used directly to elicited them

(Becker-Asano & Wachsmuth, 2010). In the case of Becker-Asano and Wachsmuth (2010), mood congruency of secondary emotions is also considered. Other proposals model the mood–emotion relationship as a modulation factor, like the case of Boukricha and Wachsmuth (2011) that implement an empathizer-mood component where a negative mood increases the potential of eliciting a negative empathic emotion and decreases the potential of a positive one. In contrast, a positive mood increases the potential of a positive empathic emotion and decreases the potential of a negative one. Proposals that do not follow any theory this relationship is modeled through the decay factor.

As a conclusion, the consideration of mood as an independent state, modeled with an appropriated theoretical base that ties it to the emotional process, guarantees a mood-congruent emotions elicitation and, therefore, more believable intelligent agents. The consideration of personality is also essential.

5.4 *Personality modeling*

Personality refers to characteristic thought patterns that persist through time and situations and distinguish one person from another (Section 3.3). Personality modeling is an important aspect to take into account when modeling believable affective agents. Nevertheless, personality is only considered in 8 of the 22 reviewed proposals. All of them are BDI extensions: personality is not modeled in any BDI formalization. In Table 4, the proposals that considered personality are presented. They have been grouped depending on whether they model personality as traits (i.e., neuroticism), temperaments (i.e., lethargic) or behavioral dispositions (i.e., irritability). Half of them follow domain-independent theoretical approaches (Eysenck, 1967; Carbonell, 1980; Merenda, 1987; McCrae & John, 1992; Ortony, 2002) which means that the considered traits/temperaments/dispositions are not tied to a particular type of situation and can be used for modeling several scenarios. Proposals that do not detail any affective theory are prone to model traits/temperaments/dispositions as domain-dependent. This is the case of Jones *et al.* (2009) and Van Dyke Parunak *et al.* (2006), that model specific traits and behavioral dispositions, tied to crisis and combats simulation situations, respectively. In the other cases, such as in Becker-Asano and Wachsmuth (2010) and Boukricha and Wachsmuth (2011), personality is modeled just as a parameter value that determines how emotional the agent is.

Personality modeling is related to emotion and mood modeling. The relationship between personality and mood is quite close, as it is generally modeled based on the idea that mood changes depending on the individual personality traits. In this regard, in most of the proposals, the agent's default or initial mood value is set according to the agent's personality traits, directly influencing its updating and decay factor over the interaction. On the contrary, Puica and Florea (2013) just model personality and emotions. In this case, the relationship is modeled through two variables, emotion intensity and decay, linked to introversion/extraversion and stable/unstable traits, respectively (Eysenck, 1967). Finally, the proposal of Neto and da Silva (2012) models the three affective aspects interconnected. In this case, the relationship is modeled through the well-known ALMA model (Gebhard, 2005) that allows mapping between emotions, mood and personality (see Section 3.3).

As a conclusion, the use of a theoretical approach allows the obtaining of domain-independent solutions and a more realistic modeling of the relationships between emotions, mood and personality. Nevertheless, these relationships need to be more intensively explored. Only the PAD and the Big Five theory have been used to model the personality–mood relationship, but as commented in Section 3, there are others interesting proposals such as Russell and PANAS theories that can be connected to Big Five, or even to another personality theory such as Eysenck (1967). The same happens with emotion–mood–personality relationship where OCC–PAD–Big Five has been the only molded through the (Gebhard, 2005) work. Thus, a formalized approach as well as a more interconnected modeling of personality in BDI agents remain to be done.

5.5 *Empathy modeling*

It is generally accepted that the modeling of the empathy in intelligent agents is very important in order to improve their human-like behavior, especially on improving human–user interaction. In spite of this,

Table 4 The proposals reviewed in this section showing the affective theory in which the mood modeling is based on and the mood states they model

Proposal	**Personality theory**	**Personality modeling/specific traits or temperament or disposition**
De Rosis *et al.* (2003)	Carbonell (1980), Ortony (2002)	Dispositions/pessimistic, optimistic
Van Dyke Parunak *et al.* (2006)	–	Dispositions/cowardice, irritability
Jones *et al.* (2009)	–	Traits/docility, curiosity, leadership, stressability, normativity
Becker-Asano and Wachsmuth (2010)	–	Temperaments/sluggish,moody
Boukricha and Wachsmuth (2011)	–	Temperaments/temperamental, lethargic
Neto and da Silva (2012)	Big Five (McCrae & John, 1992)	Traits/neuroticism, extroversion, openness, cordiality, agreeableness
Puica and Florea (2013)	Hippocrates–Galen 4 temperaments theory (Merenda, 1987) and Eysenck's theory (Eysenck, 1967)	Temperaments/sanguine, choleric, melancholic, phlegmatic
Alfonso *et al.* (2014)	Big Five (McCrae & John, 1992)	Traits/neuroticism, extroversion, openness, cordiality, agreeableness

Table 5 The reviewed proposal that model empathy, showing the empathic theory in which they are based on and the empathic emotions they model

Proposal	**Empathy theory**	**Empathic emotions modeled**
Ochs *et al.* (2010)	Own model of emotions, based on theoretical approach (Scherer, 1988; Ortony *et al.*, 1990; Lazarus, 1991; Frijda *et al.*, 1992)	Satisfaction, frustration, irritation, sadness and anger as empathic emotions
Boukricha and Wachsmuth (2011)	(Ortony *et al.*, 1990; Davis, 1994; Hoffman, 2000; De Vignemont & Singer, 2006)	Primary emotions (happy, surprised, angry, annoyed, bored) and empathic emotions related

most of the works just model empathy through the OCC emotional theory or based on simple empirical rules, as explained in Section 3.4. In this section, we will comment the two BDI extensions that were found that deal with empathy modeling (see Table 5).

Ochs *et al.* (2010) are based on empirical and theoretical approaches (Scherer, 1988; Ortony *et al.*, 1990; Lazarus, 1991; Frijda *et al.*, 1992). They make use of their own emotional model where the emotion generation process is related to a particular event and the emotional intensity corresponds to the emotional impact of the event. For modeling empathy, the intelligent agent has a mental representation of other's

Table 6 The reviewed proposals that model emotional regulation, the regulation theory they apply and the regulation strategies modeled

Proposal	**Regulation theory**	**Regulation type (ER or C)/Strategies**
De Rosis *et al.* (2003)	–	ER/Masking emotions mechanism
Bosse and Lange (2008)	Gross (1998)	ER/ A parameter that measures the emotional regulation speed to its original or default state
Van Straalen *et al.* (2009)	(Kuübler-Ross 2009)	Coping/denial, isolation, anger, bargaining, depression and acceptance
(Becker-Asano and Wachsmuth, 2010)	–	C/ Simple mechanism related to emotional reappraisal process
(Dastani & Lorini, 2012)	Lazarus (1991)	C/ related to beliefs, goals and intentions
Alfonso *et al.* (2014)	Castelfranchi (2000)	C/ Plans prioritization strategy

ER = Emotional regulation; C = Coping

mental state. The agent deduces user's intentions from the enunciation of user's communicative acts. Agent's empathic emotions elicitation depends on the success or failure of user's intentions: when events that affect user's intention occur, the agent feels an empathic emotion.

Boukricha and Wachsmuth (2011) present a more complete mechanism to model intelligent agents' empathic emotions. The emotional state and empathic emotions of the agent are represented in the PAD space, and primary and secondary emotions are modeled. The proposal is based on theoretical approaches (Ortony *et al.*, 1990; Davis, 1994; Hoffman, 2000; De Vignemont & Singer, 2006) and presents three processes to model empathy: Empathy Mechanism, Empathy Modulation and Expression of Empathy. The Empathy Mechanism is responsible for the outcome of the empathic emotions based on perceived other's emotional facial expressions. They define a predefined saliency threshold that represents agent's responsiveness to the other's situation. If the elicitation of the empathic emotion takes place, it is affirmed as belief and the second and third process occurs. The second process is the modulation of the empathic emotion intensity (Empathy Modulation component) through several predefined factors: mood, personality and social relationship to the other. Finally, the third process (Expression of Empathy) occurs and the empathic emotion expression is triggered.

As a conclusion, empathy modeling is scarcely modeled in BDI extensions and not modeled at all in formalizations: it remains an open research question.

5.6 Regulation modeling

As mentioned before (Section 3.5), there are two concepts related to regulation: coping, directed to negative emotions regulation, and emotional regulation, oriented to both positive and negative emotions. Following that distinction, the reviewed proposals are presented in Table 6. Most of the review proposals in this section are BDI extensions, but also two formalizations can be found: one related to coping, Dastani and Lorini (2012), and the other, Bosse and Lange (2008), related to emotional regulation.

Regarding coping, the proposal of Becker-Asano and Wachsmuth (2010) presents a simple coping mechanism, not based on any theory, that just considers reappraisal in the emotion module. The other three proposals deal with coping in different ways as they are based on different theories. Van Straalen *et al.* (2009) model coping in bad news scenarios. They propose a categorization of coping strategies based on the theory of Kuübler-Ross (2009) about terminal patients: Denial and Isolation, Anger, Bargaining, Depression, and Acceptance. The emotional BDI proposed by Alfonso *et al.* (2014) is based on the ideas of Castelfranchi (2000); they propose a coping component that determines if the changes on current mood deserve to take actions in cognitive processes. In those cases, the way intentions are selected is changed using a plan prioritization strategy. In the formalization proposed by Dastani and Lorini

(2012), coping modeling is based on Lazarus (1991) theory. Three coping strategies are modeled: strategies affecting intentions (removing/generating them), strategies affecting beliefs (increasing/reducing their strength) and strategies affecting goals (increasing/reducing their strength).

Regarding emotional regulation, juts two proposals have been found. De Rosis *et al.* (2003) present a simple mechanism of masking emotions formalized with context-dependent rules without explicit theoretical foundation. They consider aspects such as the interlocutors' personalities, their relationship and the social interaction context. In the case of Bosse and Lange (2008), the regulation process is modeled following Gross (1998) theory. They introduce a parameter that measures the speed to return the ERL to its default state. For example, after an unpleasant event, an agent may stay in an emotional state (i.e., angry) for a long time or leave the event behind quickly, depending on the values of the regulation parameter.

As we see, coping has been modeled far more than emotional regulation. The latter has only been modeled in very simple ways that represent a starting point in this unexplored line of investigation.

5.7 Affective influence on cognitive process

As commented in Section 3.4, nowadays, it is clear that the individual affective state affects the cognitive process. Due to this connection, it is important to consider the affective influence on the cognitive process when modeling intelligent agents.

Only three of the proposals reviewed Adam *et al.* (2009), Steunebrink *et al.* (2012) and Gluz and Jaques (2014) do not consider the affective influence on the cognitive process. All of them are based on the OCC theory that does not describe this emotional influence. The other works implement the affective influence on different steps of the BDI cognitive process. In Table 7, the emotional BDI proposals and the affective influence on the BDI cognitive process are shown.

The first column shows the theoretical base of the affective influence management. In half of them, the modeling of the affective influence is not based on theoretical foundations but just on intuitive evidences learned from psychological literature. On the contrary, others are based on affective theories well known in intelligent agent modeling such as Castelfranchi and Lorini (2003), Damasio (1994), Lazarus (1991), Loewenstein and Lerner (2003) and Frijda (1987).

As it was explained in Section 3.6, the affective influence can be modeled in several steps of the cognitive process such as beliefs, perceptions, decision-making, memory and behavior. In emotional BDIs, this is done following the steps of the BDI cognitive process (Section 4): perception, belief updating, options generation, intention filtering and actuation. In the proposals reviewed, intention filtering is the most influenced BDI cognitive process. This influence is mainly done by emotions, but in two cases mood and coping alter the process. Another cognitive aspect in which affective influence has been widely modeled is belief updating. In this case, emotions is also the affective aspect that influences the most, but mood, personality and affective capacities (coping and empathy) influences have also been modeled. The other cognitive aspect widely considered in the literature is the motivational state of the agent. In the case of the BDI architecture, this corresponds to the options generation process where the agent's desires/goals are generated. Even though it is true that emotional influence is also the most considered aspect, this is the only step where the influence of all the affective aspects has been modeled, including affective capacities such as empathy, emotional regulation and coping. Perception (influenced by emotions and mood) and actuation (influenced by emotions, mood and affective regulation) are the processes in which the affective influence has been less considered.

As it has been already commented, emotional influence is the most widely aspect considered in the literature. It is modeled by directly affecting the way perception is achieved, by directing attention toward belief-relevant information, by prioritizing desires and/or by helping to select intentions. In our review, the proposals of Pereira *et al.* (2005), Jiang *et al.* (2007) and Lejmi-Riahi *et al.* (2014) stand out. They are the most referenced ones that consider the emotional influence on most of the BDI process. Pereira *et al.* (2005) take into account the emotional influence on all the cognitive process. Nevertheless, they do not explicitly detail how the influence is implemented: they make an overview on how the emotional state of the agent modulates the decision-making process according to the availability of resources

Table 7 The proposals reviewed and the influence of agent's emotions (E) or (Other's) emotion, mood (M), personality(P), empathy (Emp), coping (C) and emotional regulation (ER) on the BDI cognitive processes, as well as the theoretical bases of these influences

Proposal	Theoretical basses	Perception	Belief updating	Options generation	Intensions filtering	Actuation
De Rosis *et al.* (2003)	–	–	–	E/P	–	E/ER
Hernández *et al.* (2004)	Castelfranchi (2000), Canamero (2000)	–	E	E	–	–
Pereira *et al.* (2005)	–	E	E	E	E	E
Van Dyke Parunak *et al.* (2006)	–	–	–	–	E	–
Meyer (2006)	–	–	–	–	E	–
Jiang *et al.* (2007)	Damasio (1994), Frijda *et al.* (2000b)	–	E	E	E	–
Bosse and de Lange (2008)	Gross (1998)	–	–	E	–	–
Jones *et al.* (2009)	–	–	E/P	P	E	–
Bosse and Zwanenburg (2009)	Castelfranchi and Lorini (2003)	–	–	–	–	E
Van Straalen *et al.* (2009)	Folkman and Lazarus (1980), Carver *et al.* (1989)	–	–	–	–	C
Becker-Asano and Wachsmuth (2010)	–	–	–	–	E	–
Ochs *et al.* (2010)	–	Other's E	–	Emp	–	–
Boukricha and Wachsmuth (2011)	–	Other's E	Emp	–	–	–
Bosse and Höhle (2011)	–	–	E	E/M	E	M
Neto and da Silva (2012)	Gazzaniga and Heatherton (2006), Izard (1993), Damasio (1994)	M	M	–	M	–
Dastani and Lorini (2012)	Lazarus (1991)	–	C	C	E/C	–
Puica and Florea (2013)	–	E	E	–	E	–
Lejmi-Riahi *et al.* (2014)	Loewenstein and Lerner (2003)	E	E	E	E	–
Alfonso *et al.* (2014)	Niedenthal and Setterlund (1994), Castelfranchi (2000)	M	–	–	M/C	–

and capabilities (see Section 5.1). In the case of Jiang *et al.* (2007), they take into account the influence of primary and secondary emotions on belief updating, option generation and intention filtering. They use emotions to set the priority of desires and help to decide intentions. To do so, they follow the idea of some researchers (Camerer *et al.*, 2003) that point out that decision-making may actually be driven by emotions, since people with emotional impairments have trouble making decisions. The proposal of Lejmi-Riahi *et al.* (2014) takes into account the influence of secondary emotions in almost all the BDI cognitive processes. They follow the theory of Loewenstein and Lerner (2003) that establishes two ways in which emotions can influence at the time of decision-making: as anticipated consequences (expected emotions) and as feelings experienced at the time of decision-making (immediate emotions). Immediate emotions influence all cognitive processes except actuation. In the case of expected emotions, the influence is modeled in the options generation process. This way, they propose a complex cognitive-affective processing that makes it possible to model other types of behaviors that are not directly product of decision.

Another way to influence decision-making is through a more permanent affective state such as mood. Only three of the reviewed proposals consider this influence. In the case of Bosse and Höhle (2011) (i.e., the one that presents a formalization), they do not detail the theoretical foundation of their modeling. Mood and emotional influence are modeled in the same way: as rules that are mostly domain-specific and define the impact of emotions and mood on the BDI cognitive process; for instance, a positive mood increases the desire to cooperate with teammates and a negative mood increases the desire to behave aggressively. The other two proposals model mood's influence independently of emotional influence. Their mood's influence modeling is based on theoretical foundations and focuses on perception and decision-making (and in memory in the case of Neto and da Silva (2012)). Both proposals model the influence on perception based on the same idea: mood can intensify or blur perceptions: depending on agent's mood, the same sensation may generate different perceptions. Regarding the mood's influence on decision-making, they model it in different ways. In Neto and da Silva (2012) proposal, it is modeled in the updating beliefs and intention filtering processes. This is implemented through emotional markers (Damasio, 1994) that allow tagging actions with emotional states, in such a way that actions can be selected through the corresponding emotional states that tag them, and vice versa. In contrast, Alfonso *et al.* (2014) only model mood influence on the intention filtering process: if the changes experimented in the current mood deserve to take actions, this will determine the intentions selected.

Regarding personality, there are two proposals (BDI extensions) that model its influence on the BDI cognitive process: De Rosis *et al.* (2003) and Jones *et al.* (2009). In both proposals, the personality influences the options generation process, but in different ways. In the case of De Rosis *et al.* (2003), personality traits are related to certain goal weights directly affecting the intensity of the elicited emotions. In Jones *et al.* (2009), they also consider the influence of the agent's physiological state but it is not detailed. They also model the influence of personality (and emotions) on the belief updating process, in particular, in the way new beliefs are interpreted.

Affective capabilities also influence the BDI cognitive processes. As in precedent sections, we will analyze affective regulation and empathy.

Proposals that consider regulation (coping and emotional regulation) influence have all a strong theoretical foundation. In general, both emotional regulation and coping have direct influence on agent decision-making and final behavior. Coping modeling has been modeled as a mechanism that directly affects the intention filtering and/or actuation process, determining various types of responses. In Alfonso *et al.* (2014), coping influences the intention filtering process through a mechanism of plan's prioritization according to the agent's current mood. In the case of Van Straalen *et al.* (2009), as their particularly interested in the way the agent responds to bad news, coping alters the actuation process, both influencing the behavior selection and the manifestation of the selected behavior. Dastani and Lorini (2012) present a more complete model. Following the idea that coping can be seen as a cognitive mechanism whose aim is to discharge a certain emotion by modifying one or more of the mental attitudes that triggered the emotion, they consider it in beliefs updating, options generation and intention filtering. This is done by looking over an intention, by changing the agent's interpretation of the situation, by changing a belief or a goal's strength.

Regarding emotional regulation influence, just one proposal has been found that models it. In De Rosis *et al.* (2003), the influence of emotional regulation is in the actuation process: it affects directly agent's expression, masking elicited emotions. This does not necessarily imply the elimination of an emotion but the variation of its intensity or its duration. In the case of Bosse and Lange (2008), that also consider emotional regulation, they do not model a direct influence of regulation on the agent's cognitive process. The influence is indirect on the agent ERL, which in turn influences the motivational state of the agent.

The influence of empathy in the cognitive process is modeled in the two proposals that consider empathy modeling (Section 3.5). In both of them, the other's emotions influence the cognitive process in the perception process. Besides, the agent empathic emotions influence the option generation process in the case of Ochs *et al.* (2010) and the belief updating in Boukricha and Wachsmuth (2011) proposal.

As a conclusion, the BDI cognitive phases influenced and how this influence is modeled varies greatly from one proposal to another. Emotions influence is taken into account in most of the proposals, mood and personality to a lesser extent. The belief updating process and the intention filtering process are the processes most influenced by emotions, mood and, in a lesser extent, by personality. Options generation process is frequently influenced by a combination of some affective aspects such as emotions and mood or emotions and personality. The influence of affective capacities (affective regulation and empathy) is scarcely modeled and remains nowadays an open line of research. In general, the actuation process is the most influenced by coping and emotional regulation as it is related to the final behavior of the agent. Empathy has been modeled to influence the belief processing and the options generation processes.

6 Modeling affect in BDI agents: good practices and open questions

After the review presented in the last section, some common features, good practices and open questions can be extracted. We present them grouped by affective aspect.

6.1 Emotions

Emotions are short-lived affective states that arise suddenly in the context of a specific stimulus. Nowadays, it is recognized as mandatory to endow the agent with an emotional model if believability is sought. In most of the proposals, emotions influence the cognitive process in several ways: in perception, motivation and decision-making processes or at least a combination of them.

6.1.1 Good practices

- Emotional modeling should be based on an established emotional theory. It is recommended that emotional modeling includes both primary and secondary emotions.
- It is recommended to combine emotional elicitation theories with dimensional theories to allow a more complete emotional modeling and to facilitate the modeling of emotion–mood dynamics.
- It is important to quantify emotions in terms of their intensity, to be able to differentiate between emotions of the same type such as irritation, anger and rage. It makes it possible also to consider if an emotion is actually experienced once elicited.
- Emotions should influence perceived information and all processes of BDI reasoning (updating beliefs, generating desires and obtaining intentions). To do so, several theories of emotion have to be combined.

6.1.2 Open questions

- It is still missing a proposal that allows managing the emotional influence on all the BDI phases.
- Tertiary emotions' modeling is missing.
- It is still a challenge an efficient modeling of mood-congruent emotion elicitation.

6.2 Mood

Mood, as a more durable affective state, has been shown to impact a range of cognitive, perceptual and behavioral processes, and the emotional elicitation (mood-congruent emotions). Nevertheless, in many of the proposals, it is just considered as an unattributed affective basis to which emotions decay.

6.2.1 Good practices

- Mood should be modeled as an independent affective state rather than just as the decay of emotions.
- Mood should influence all the phases of the BDI process.
- The relationship between mood and other affective states such as emotions and personality should be considered:
  - It is common to consider personality when calculating the agent's initial mood or mood 'by default'.
  - Mood-updating process should be influenced be the elicited emotions. A good practice is to relate it to the intensities of the emotions experienced.
  - The mood of the agent should be taken into account in the emotion generation, in particular, in the calculation of the intensity of the elicited emotions.

6.2.2 Open questions

- It is still missing a proposal that models in efficient way the mood's influence on the emotion elicitation process to obtain mood-congruent emotions.
- The modeling of the influence of mood on every BDI cognitive phase is still missing.

6.3 Personality

Personality is considered a long-term affective state that does not change over time. This way personality can be used to differentiate one agent from another.

6.3.1 Good practices

- Even though the Big Five model is the most used personality model, others such as The four temperaments and the one proposed by Carbonell (1980) should be also considered depending on the agentś context.
- Personality should be taken into account on the emotional generation process and mood decay time.
- Personality should influence the BDI cognitive process together to other affective states.

6.3.2 Open questions

- It is still missing an efficient modeling of personality's influence on cognitive and affective processes.

6.4 Empathy

Empathy is the capacity to understand or feel what another person is experiencing from within his or her frame of reference. Empathic experience is related to the other's cognitive-affective state. Empathy should be considered if believable agent–user interaction is sought.

6.4.1 Good practices

- To model empathy, it is necessary to have the representation of other's mental attitudes.
- Empathy should be modeled in its two variants: cognitive and affective.
- Empathic emotions should influence the decision process as do non-empathic ones.

6.4.2 Open questions

- A complete and theoretically based modeling of empathy in emotional BDI agents is still missing.

6.5 Regulation

Regulation modeling is focused on coping with stressing situations and regulating both negative and positive emotions. In the proposals that model coping, the regulatory process influences the selection of the agent's behavior and how it is expressed. On the other hand, the proposals that model emotional regulation are focused on the regulation of the emotional response.

6.5.1 Good practices

- The regulation process should influence not only the emotions displayed by the agent but also its behavior.
- Both coping and emotional regulation should be implemented. To do so, the Gross theory (1998) stands out as it allows coping strategies to be integrated into the emotional regulatory process.

6.5.2 Open questions

- The works that consider regulatory mechanisms are scarce and it is still an open question how to integrate these issues into the general BDI scheme.
- EER aiming to regulate emotion both within the provider and across interaction partners has not been yet considered in any proposal.
- Mood regulation modeling is missing. Larsen theory (Larsen, 2000) could be a good starting point. The consideration of mood regulation would open the door to a more wide consideration of regulation (the so-called affective regulation).

As it can be seen, in spite of the great advances in the modeling of affective intelligent agents, there is no proposal that models the influence of emotions, mood and personality in the all BDI phases. In this regard, it is interesting to highlight, that in some of the works, the affective state of the agent is a component composed by personality, mood and emotions, and it is the result of this interaction what influences the cognitive BDI process. This is an interesting point of view that allows the modeling of emotion–mood–personality dynamics and its influence on BDI cognitive process. Other relevant affective capabilities such as regulation and empathy still require a strong research effort. It is true that in many cases the open research questions detected are related to open psychological questions without generally accepted models. In this regard, a more intense interdisciplinary exchange between psychology and computer science could make it possible to significantly advance in the challenge of understanding and modeling human emotions.

7 Conclusions

In the last decades, the BDI schema has been extended to incorporate affective states and affective processes such as empathy, emotional regulation and coping. First, the main theories applied to the modeling of affective states and affective capabilities in intelligent agents have been revised. Second, an exhaustive review of affective BDI proposals has been presented. The revised works have been first classified according to their nature as formalizations or BDI architecture extensions. Then, they have been classified and analyzed according to the affective aspects they consider. From the review, it can be seen that, nowadays, the tendency is to deal not only with short-term affective states such as emotions but to incorporate mid- and long-term ones such as mood and personality traits. Besides, works try not just to add but to integrate affective states in all phases of the BDI cognitive schema. Affective processes such as empathy or regulation are also beginning to be seriously considered, but their implementation still remains a challenging task. In fact, and in spite of the great effort done in the last years to incorporate affect into BDI agents, several open questions remain and a complete and integrated emotional BDI proposal is still missing.

Acknowledgments

This work has been partly financed by the Aragonese Government and the UE through the FEDER 2014-2020 'Construyendo Europa desde Aragón' action (Group T25_17D) and Banco Santander and Universidad de Zaragoza through a mobility grant for Latin American students. We want to thank psychologists Antonio Aguelo and Teresa Coma for the fruitful discussions about affect.

References

Adam, C., Gaudou, B., Herzig, A. & Longin, D. 2006. OCC's emotions: a formalization in a BDI logic. *In International Conference on Artificial Intelligence: Methodology, Systems, and Applications*, 24–32. Springer Berlin Heidelberg.

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

Alfonso, B., Vivancos, E. & Botti, V. J. 2014. An open architecture for affective traits in a BDI agent. In *Proceedings of the 6th ECTA. Part of the 6th IJCCI*, 320–325. SCITEPRESS (Science and Technology Publications, Lda).

Anderson, J. R. 1993. *Rules of the Mind*. Lawrence Erlbaum Associates, Inc.

André, E., Klesen, M., Gebhard, P., Allen, S. & Rist, T. 2000. Integrating models of personality and emotions into lifelike characters. In *Affective Interactions*, 150–165. Springer Berlin Heidelberg.

Angie, A. D., Connelly, S., Waples, E. P. & Kligyte, V. 2011. The influence of discrete emotions on judgement and decision-making: a meta-analytic review. *Cognition & Emotion* **25**(8), 1393–1422.

Becker-Asano, C. 2014. WASABI for affect simulation in human-computer interaction. In *Proceeding International Workshop on Emotion Representations and Modelling for HCI Systems*.

Becker-Asano, C. & Wachsmuth, I. 2008. Affect simulation with primary and secondary emotions. *International Workshop on Intelligent*.

Becker-Asano, C. & Wachsmuth, I. 2010. Affective computing with primary and secondary emotions in a virtual human. *Autonomous Agents and Multi-Agent Systems* **20**(20), 32–49.

Bevacqua, E., Mancini, M. & Pelachaud, C. 2008. A listening agent exhibiting variable behaviour. In *International Workshop on Intelligent Virtual Agents*. Springer, Berlin, Heidelberg, 262–269.

Bevacqua, E., Prepin, K., Niewiadomski, R., de Sevin, E. & Pelachaud, C. 2010. Greta: towards an interactive conversational virtual companion. *Artificial Companions in Society: Perspectives on the Present and Future* 143–156.

Bickmore, T. W., Utami, D., Matsuyama, R. & Paasche-Orlow, M. K. 2016. Improving access to online health information with conversational agents: a randomized controlled experiment. *Journal of Medical Internet Research* **18**(1), e1.

Blanchette, I. & Richards, A. 2010. The influence of affect on higher level cognition: a review of research on interpretation, judgement, decision making and reasoning. *Cognition & Emotion* **24**(4), 561–595.

Bosse, T. & Höhle, D. 2011. Enhancing believability of virtual soccer players: application of a BDI-model with emotions and trust. In *Developing Concepts in Applied Intelligence*, Mehrotra, K. G., Mohan, C. O., Jae, C., Varshney, P. K. & Ali, M. (eds). Springer Berlin Heidelberg, 119–128.

Bosse, T., Jonker, C. M., Van Der Meij, L. & Treur, J. 2007a. A language and environment for analysis of dynamics by simulation. *International Journal on Artificial Intelligence Tools* **16**(3), 435–464.

Bosse, T. & de Lange, F. P. J. 2008. Development of virtual agents with a theory of emotion regulation. In *Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, 2, 461–468. IEEE Computer Society.

Bosse, T., Pontier, M. & Treur, J. 2007b. A computational model for adaptive emotion regulation. In *Proceedings of the 8th International Conference on Cognitive Modeling, ICCM*, 7, 187–192. Taylor and Francis/Psychology Press.

Bosse, T. & Zwanenburg, E. 2009. There's always hope: enhancing agent believability through expectation-based emotions. In *Proceedings – 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, ACII 2009*, 1–8. IEEE.

Boukricha, H. & Wachsmuth, I. 2011. Empathy-based emotional alignment for a virtual human: a three-step approach. *KI – Künstliche Intelligenz* **25**(3), 195–204.

Bratman, M. 1987. *Intention, Plans, and Practical Reason*. Harvard University Press.

Braubach, L., Lamersdorf, W. & Pokahr, A. 2003. Jadex: implementing a BDI-infrastructure for JADE. *EXP in Search of Innovation* **3**(3), 76–85.

Breazeal, C. 2002. *Designing Sociable Robots*. MIT Press.

Breazeal, C. 2003. Emotion and sociable humanoid robots. *International Journal of Human–Computer Studies* **59**(1), 119–155.

Broekens, J., DeGroot, D. & Kosters, W. A. 2008. Formal models of appraisal: theory, specification, and computational model. *Cognitive Systems Research* **9**(3), 173–197.

Camerer, C. F., Loewenstein, G. & Rabin, M. 2003. *Advances in Behavioral Economics*. Princeton University Press.

Canamero, D. 2000. *Designing Emotions for Activity Selection*. Technical report, University of Aarhus, Denmark.

Carbonell, J. G. 1980. Towards a process model of human personality traits. *Artificial Intelligence* **15**(1), 49–74.

Carver, C. S., Scheier, M. F. & Weintraub, J. K. 1989. Assessing coping strategies: a theoretically based approach. *Journal of Personality and Social Psychology* **56**(2), 267.

Castelfranchi, C. 2000. Affective appraisal versus cognitive evaluation in social emotions and interactions. In *Affective Interactions*, Paiva, A. (ed.). Springer Berlin Heidelberg, 70–106.

Castelfranchi, C. & Lorini, E. 2003. Cognitive anatomy and functions of expectations. In *IJCAI03 Workshop on Cognitive Modeling of Agents and Multi-Agent Interactions*, 9–11. Morgan Kaufmann Publishers.

Cochran, R. E. 2006. Modeling emotion: arousal's impact on memory. In *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, 1133–1138. Erlbaum.

Cohen, P. R. & Levesque, H. J. 1990. Intention is choice with commitment. *Artificial Intelligence* **42**(2–3), 213–261.

Corchado, J. M., Pavón, E. S., Corchado, E. S. & Castillo, L. F. 2004. Development of CBR-BDI agents: a tourist guide application. In *European Conference on Case-based Reasoning*, 547–559. Springer Berlin Heidelberg.

Damasio, A. 1994. *Descartes' Error: Emotion, Reason, and the Human Brain*. Putnam (Grosset Books).

Damasio, A. R. 2005. *Descartes' Error: Emotion, Reason, and the Human Brain*. Penguin Group US.

Dastani, M. & Lorini, E. 2012. A logic of emotions: from appraisal to coping. In *11th International Conference on Autonomous Agents and Multiagent Systems*, 1133–1140. International Foundation for Autonomous Agents and Multiagent Systems.

Davis, M. H. 1994. *Empathy: A Social Psychological Approach*. Westview Press.

Desmet, P. 2015. Design for mood: twenty activity-based opportunities to design for mood regulation design. *International Journal of Design* **9**(2), 1–19.

Dias, J., Mascarenhas, S. & Paiva, A. 2014. Fatima modular: towards an agent architecture with a generic appraisal framework. In *Emotion Modeling*, Bosse, T., Broekens, J., Dias, J. & van der Zwaan, J. M. (eds). Springer International Publishing, 44–56.

D'Inverno, M., Luck, M., Georgeff, M., Kinny, D. & Wooldridge, M. 2004. The DMARS architecture: a specification of the distributed multi-agent reasoning system. *Autonomous Agents and Multi-Agent Systems* **9**(1–2), 5–53.

Dolan, R. J. 2002. Emotion, cognition, and behavior. *Science* **298**(5596), 1191–1194.

Durupinar, F., Allbeck, J. M., Pelechano, N. & Badler, N. I. 2008. Creating crowd variation with the OCEAN personality model. In *7th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2008)*, 3.

Van Dyke Parunak, H., Bisson, R., Brueckner, S., Matthews, R. & Sauter, J. 2006. A model of emotions for situated agents. In *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems*, 993–995. ACM.

Ekman, P., Friesen, W. V. & Ellsworth, P. 1972. *Emotion in the Human Face: Guidelines for Research and an Integration of Findings*. Elsevier.

El-Nasr, M. S., Yen, J. & Ioerger, T. R. 2000. Flame—fuzzy logic adaptive model of emotions. *Autonomous Agents and Multi-Agent Systems* **3**(3), 219–257.

Emonts, M., Row, R., Johnson, W. L., Thomson, E., De Silva Joyce, H., Gorman, L. G. & Carpenter, R. 2012. Integration of social simulations into a task-based blended training curriculum. In *Land Warfare Conference*.

Eysenck, H. J. 1967. *The Biological Basis of Personality*, **689**. Transaction Publishers.

Eysenck, H. J. 1991. Dimensions of personality: 16, 5 or 3?—criteria for a taxonomic paradigm. *Personality and Individual Differences* **12**(8), 773–790.

Folkman, S. & Lazarus, R. S. 1980. An analysis of coping in a middle-aged community sample. *Journal of Health and Social Behavior* **21**(4), 219–239.

Frijda, N. H. 1986. *The Emotions*. Cambridge University Press.

Frijda, N. H. 1987. Emotion, cognitive structure, and action tendency. *Cognition & Emotion* **1**(2), 115–143.

Frijda, N. H., Manstead, A. S. R. & Bem, S. 2000a. The influence of emotions on beliefs. In *Emotions and Beliefs: How Feelings Influence Thoughts*, Pergamon, 1–9.

Frijda, N. H., Manstead, A. S. R. & Bem, S. (eds) 2000b. *Emotions and Beliefs: How Feelings Influence Thoughts*. Cambridge University Press.

Frijda, N. H., Ortony, A., Sonnemans, J. & Clore, G. L. 1992. The complexity of intensity: issues concerning the structure of emotion intensity. In *Review of Personality and Social Psychology, No. 13. Emotion*, Clark, M. S. (ed.). SAGE Publications, 60–89.

Fum, D. & Stocco, A. 2004. Memory, emotion, and rationality: an ACT-R interpretation for gambling task results. In *Proceedings of the 6th International Conference on Cognitive Modeling*, 106–111. Lawrence Erlbaum Associates.

Gazzaniga, M. S. & Heatherton, T. F. 2006. *Psychological Sciences: Mind, Brain, and Behavior*, W.W. Norton.

Gebhard, P. 2005. ALMA – a layered model of affect. In *Proceedings of the 4th International Joint Conference on Autonomous Agents and Multiagent Systemsent Systems*, 29–36. ACM.

Georgeff, M. P. & Lansky, A. L. 1987. Reactive reasoning and planning. In *Proceedings of the Sixth National Conference on Artificial Intelligence (AAAI-87)*, 87, 677–682.

Gibson, L. 2006. Mirrored emotion. *The University of Chicago Magazine* **94**(4), 1–9.

Glore, G. L. & Gasper, K. 2000. Feeling is believing: some affective influences on belief. In *Emotions and Beliefs: How Feelings Influence Thoughts*, Frijda, N. H., Manstead, A. S. R. & Bem, S. (eds). Cambridge University Press, 10–44.

Gluz, J. & Jaques, P. 2014. A probabilistic implementation of emotional BDI agents. In *Proceedings of the 6th International Conference on Agents and Artificial Intelligence*, 121–129. SCITEPRESS – Science and and Technology Publications.

Gluz, J. & Jaques, P. A. 2017. A probabilistic formalization of the appraisal for the OCC event-based emotions. *Journal of Artificial Intelligence Research* **58**(1), 627–664.

Gratch, J. & Marsella, S. 2004. A domain-independent framework for modeling emotion. *Cognitive Systems Research* **5**(4), 269–306.

Green, O. H. & Green, O. H. 1992. *The Emotions: A Philosophical Theory*, **53**. Springer Science & Business Media.

Gris, I., Rivera, D. A., Rayon, A., Camacho, A. & Novick, D. 2016. Young merlin: an embodied conversational agent in virtual reality. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction – ICMI 2016*, 425–426. ACM Press.

Gross, J. J. 1998. The emerging field of emotion regulation: an integrative review. *Review of General Psychology* **2**(5), 271–299.

Gunes, H., Schuller, B., Pantic, M. & Cowie, R. 2011. Emotion representation, analysis and synthesis in continuous space: a survey. In *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference*, 817–834. IEEE.

Hamburg, M. E., Finkenauer, C. & Schuengel, C. 2014. Food for love: the role of food offering in empathic emotion regulation. *Frontiers in Psychology* **5**(32).

Hernández, D. J., Déniz, O., Lorenzo, J. & Hernández, M. 2004. BDIE: a BDI like architecture with emotional capabilities. In *American Association for Artificial Intelligence Spring Symposium*.

Herzig, A. & Longin, D. 2004. C&L intention revisited. In *Principles of Knowledge Representation and Reasoning – KR*, 527–535.

Hindriks, K. V., De Boer, F. S., Van der Hoek, W. & Meyer, J.-J. Ch. 1999. Agent programming in 3APL. *Autonomous Agents and Multi-Agent Systems* **2**(4), 357–401.

Hoffman, M. L. 2000. *Empathy and Moral Development*. Cambridge University Press.

Huber, M. J. 1999. JAM: a BDI-theoretic mobile agent architecture. In *Proceedings of the third annual conference on Autonomous Agents*, 236–243. ACM.

Hudlicka, E. 1998. Modeling emotion within symbolic cognitive architectures. *Technical Report FS-98-03, Papers from the 1998 AAAI Fall Symposium*.

Hudlicka, E. 2004. *Two Sides of Appraisal: Implementing Appraisal and Its Consequences within a Cognitive Architecture*. AAAI Press.

Hudlicka, E. 2008. What are we modeling when we model emotion? In *AAAI Spring Symposium: Emotion, Personality, and Social Behavior*, 8.

Hudlicka, E. 2011. Guidelines for designing computational models of emotions. *International Journal of Synthetic Emotions* **2**(1), 26–79.

Hughes, G. E. & Cresswell, M. J. 1996. *A New Introduction to Modal Logic*. Psychology Press.

Ingrand, F. F., Rao, A. S. & Georgeff, M. P. 1992. An architecture for real-time reasoning and system control. *IEEE Expert* **7**(6), 33–44.

Izard, C. E. 1993. Four systems for emotion activation: cognitive and noncognitive processes. *Psychological Review* **100**(1), 68–90.

Jeon, M. 2017. *Emotions and Affect in Human Factors and Human–Computer Interaction*. Academic Press.

Jiang, H., Vidal, J. M. & Huhns, M. N. 2007. EBDI: an architecture for emotional agents. In *Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems*, 5, 11. ACM.

Jones, H., Saunier, J. & Lourdeaux, D. 2009. Personality, emotions and physiology in a BDI agent architecture: the PEP→BDI model. In *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, 2, 263–266. IEEE Computer Society.

Kleinginna Jr, P. R. & Kleinginna, A. M. 1981. A categorized list of emotion definitions, with suggestions for a consensual definition. *Motivation and Emotion* **5**(4), 345–379.

Kopp, S., Gesellensetter, L., Krämer, N. C. & Wachsmuth, I. 2005. A conversational agent as museum guide – design and evaluation of a real-world application. In *International Workshop on Intelligent Virtual Agents*, 329–343. Springer Berlin Heidelberg.

Kuübler-Ross, E. 2009. *On Death and Dying: What the Dying Have to Teach Doctors, Nurses, Clergy and Their Own Families*. Taylor & Francis.

Laird, J. E., Newell, A. & Rosenbloom, P. S. 1987. Soar: an architecture for general intelligence. *Artificial Intelligence* **33**(1), 1–64.

Larsen, R. J. 2000. Toward a science of mood regulation. *Psychological Inquiry* **11**(3), 129–141.

Larsen, R. J. and Ketelaar, T. 1989. Extraversion, neuroticism and susceptibility to positive and negative mood induction procedures. *Personality and Individual Differences* **10**(12), 1221–1228.

Larsen, R. J. & Ketelaar, T. 1991. Personality and susceptibility to positive and negative emotional states. *Journal of Personality and Social Psychology* **61**(1), 132–140.

Lazarus, R. S. 1991. *Emotion and Adaptation*. Oxford University Press.

Lazarus, R. S. & Folkman, S. 1984. *Stress, Appraisal, and Coping*. Springer Publishing Co.

LeDoux, J. E. 1996. *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. Simon & Schuster.

Lejmi-Riahi, H., Kebair, F. & Said, L. B. 2014. Agent decision-making under uncertainty: towards a new E-BDI agent architecture based on immediate and expected emotions. *International Journal of Computer Theory and Engineering* **6**(3), 254.

Loewenstein, G. & Lerner, J. 2003. The role of affect in decision making. *Handbook of Affective Science* **619**(642), 3.

Lönnqvist, J. E., Verkasalo, M. & Walkowitz, G. 2001. It pays to pay–big five personality influences on co-operative behaviour in an incentivized and hypothetical Prisoner's dilemma game. *Personality and Individual Differences* **50**(2), 300–304.

Marsella, S. C. & Gratch, J. 2009. EMA: a process model of appraisal dynamics. *Cognitive Systems Research* **10**(1), 70–90.

McCrae, R. R. & John, O. P. 1992. An introduction to the five-factor model and its applications. *Journal of Personality* **60**(2), 175–215.

McEachan, R. R. C., Sutton, S. & Myers, L. 2010. Mediation of personality influences on physical activity within the theory of planned behaviour. *Journal of Health Psychology* **15**(8), 1170–1180.

McQuiggan, S. W. & Lester, J. C. 2006. Learning empathy: a data-driven framework for modeling empathetic companion agents. In *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems*, 961–968, edited by ACM.

Mehrabian, A. 1996a. Analysis of the big-five personality factors in terms of the PAD temperament model. *Australian Journal of Psychology* **48**(2), 86–92.

Mehrabian, A. 1996b. Pleasure-arousal-dominance: a general framework for describing and measuring individual differences in temperament. *Current Psychology* **14**(4), 261–292.

Merenda, P. F. 1987. Toward a four-factor theory of temperament and/or personality. *Journal of Personality Assessment* **51**(3), 367–374.

Meyer, J. J. Ch. 2006. Reasoning about emotional agents. *International Journal of Intelligent Systems* **21**(6), 601–619.

Meyer, J. J. Ch. & Van Der Hoek, W. 2004. *Epistemic Logic for AI and Computer Science*. Cambridge University Press.

Meyer, J. J. Ch., Van Der Hoek, W. & van Linder, B. 1999. A logical approach to the dynamics of commitments. *Artificial Intelligence* **113**(1), 1–40.

Moore, R. C. 1980. Reasoning about knowledge and action. In *SRI International*.

Morley, D. & Myers, K. 2004. The SPARK agent framework. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems*, 2, 714–721. IEEE Computer Society.

Neto, A. F. B. & da Silva, F. S. C. 2012. A computer architecture for intelligent agents with personality and emotions. In *Human-Computer Interaction: The Agency Perspective*, 263–285. Springer Berlin Heidelberg.

Neumann, R., Seibt, B. & Strack, F. 2001. The influence of mood on the intensity of emotional responses: disentangling feeling and knowing. *Cognition & Emotion* **15**(6), 725–747.

Newell, A. 1994. *Unified Theories of Cognition*. Harvard University Press.

Niedenthal, P. M. & Setterlund, M. B. 1994. Emotion congruence in perception. *Personality and Social Psychology Bulletin* **20**(4), 401–411.

Nunamaker Jr, J. F., Derrick, D. C., Elkins, A. C., Burgoon, J. K. & Patton, M. W. 2011. Embodied conversational agent-based kiosk for automated interviewing. *Journal of Management Information Systems* **28**(1), 17–48.

Oatley, K. & Jenkins, J. M. 1996. *Understanding Emotions*. Blackwell Publishing.

Ochs, M., Paris, U., Pelachaud, C., Paris, U. & Sadek, D. 2008. An empathic virtual dialog agent to improve human-machine interaction. In *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems*, 1, 89–96. International Foundation for Autonomous Agents and Multiagent Systems.

Ochs, M., Sadek, D. & Pelachaud, C. 2010. A formal model of emotions for an empathic rational dialog agent. *International Journal of Autonomous Agents and Multi-Agent Systems* **24**(3), 410–440.

Van Oijen, J., Van Doesburg, W. & Dignum, F. 2011. Goal-based communication using BDI agents as virtual humans in training: an ontology driven dialogue system. In *Agents for Games and Simulations II, Lecture Notes in Computer Science*, Dignum, F. (ed). Springer Verlag, 38–52.

Ortony, A. 2002. On making believable emotional agents believable. In *Emotions in Humans and Artifacts*, Trappl, R., Petta, P. & Payr, S. (eds). The MIT Press, 189–211.

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

Pacherie, E. 2004. L'empathie et Ses Degrés. In *L'empathie*, sous la dir Berthoz, de A. & Jorland, G., 149–181. Editions Odile Jacob.

Paiva, A., Dias, J., Sobral, D., Woods, S. & Hall, L. 2004. Building empathic lifelike characters: the proximity factor. In *Workshop on Empathic Agents, AAMAS*, 4.

Pereira, D., Oliveira, E., Moreira, N. & Sarmento, L. 2005. Towards an architecture for emotional BDI agents. In *EPIA 2005 12th Portuguese Conference on Artificial Intelligence, Universidade da Beira Interior*, 40–46. IEEE.

Pérez, J., Cerezo, E., Serón, F. J. & Rodríguez, L.-F. 2016. A cognitive-affective architecture for ECAs. *Biologically Inspired Cognitive Architectures* **18**, 33–40.

Plutchik, R. 2001. The nature of emotions. *American Scientist* **89**(4), 344.

Pollack, M. E. & Ringuette, M. 1990. Introducing the tileworld: experimentally evaluating agent architectures. *AAAI* **90**, 183–189.

Puica, M.-A. & Florea, A.-M. 2013. Emotional belief-desire-intention agent model: previous work and proposed architecture. *International Journal of Advanced Research in Artificial Intelligence(IJARAI)* **2**(2), 1–8.

Rao, A. S. & Georgeff, M. P. 1992. An abstract architecture for rational agents. KR'92, 439–449.

Rao, A. S. 2009. AgentSpeak(L): BDI agents speak out in a logical computable language. In *European Workshop on Modelling Autonomous Agents in a Multi-Agent World*, 42–55. Springer Berlin Heidelberg.

Rao, A. S. & Georgeff, M. P. 1995. BDI agents: from theory to practice. *Proceedings of the First International Conference on Multiagent Systems (ICMAS95)* 95, 312–319.

Van Reekum, C. M. 2000. *Levels of Processing in Appraisal: Evidence from Computer Game Generated Emotions*. University of Geneva.

Reilly, W. S. 1996. *Believable Social and Emotional Agents*. PhD thesis, Carnegie Mellon University, School of Computer Science.

Reisenzein, R. 2006. Arnold's theory of emotion in historical perspective. *Cognition and Emotion* **20**(7), 920–951.

Reisenzein, R., Hudlicka, E., Dastani, M., Gratch, J., Hindriks, K., Lorini, E. & Meyer, J.-J. Ch. 2013. Computational modeling of emotion: toward improving the inter- and intradisciplinary exchange. *IEEE Transactions on Affective Computing* **4**(3), 246–266.

Reisenzein, R. & Junge, M. 2012. Language and emotion from the perspective of the computational belief-desire theory of emotion. *Dynamicity in Emotion Concepts* **27**, 37–59.

Remington, N. A., Visser, P. S., Diener, E, Barrett, L. F., Green, D., Mayer, J., Rusting, C. & Watson, D. 2000. Reexamining the circumplex model of affect. *Journal of Personality and Social Psychology* **79**(2), 286–300.

Roseman, I. J. & Smith, C. 2001. Appraisal Theory: Overview, Assumptions, Varieties, Controversies. In *Appraisal Processes in Emotion: Theory, Methods, Research*, Oxford University Press, 3–19.

Roseman, I. J., Spindel, M. S. & Jose, P. E. 1990. Appraisals of emotion-eliciting events: testing a theory of discrete emotions. *Journal of Personality and Social Psychology* **59**(5), 899–915.

Rosenbloom, P. S. 2013. The sigma cognitive architecture and system. *AISB Quarterly* **136**, 4–13.

De Rosis, F., Pelachaud, C., Poggi, I., Carofiglio, V. & De Carolis, B. 2003. From Greta's mind to her face: modelling the dynamics of affective states in a conversational embodied agent. *International Journal of Human–Computer Studies* **59**(1), 81–118.

Russell, J. A. 1980. A circumplex model of affect. *Journal of Personality and Social Psychology* **39**(6), 1161–1178.

Russell, S. J., Norvig, P., Canny, J. F., Malik, J. M. & Edwards, D. D. 2003. *Artificial Intelligence: A Modern Approach*, **2**. Prentice hall.

Scherer, K. R. 1984. On the nature and function of emotion: a component process approach. In *Approaches to Emotion*, Scherer, K. R. & Ekman, P. E. (eds). Psychology Press, **2293**, 317.

Scherer, K. R. 1988. Criteria for emotion-antecedent appraisal: a review. In *Cognitive Perspectives on Emotion and Motivation*, Hamilton V., Bower G.H. & Frijda N.H. (eds). Springer Netherlands, 89–126.

Scherer, K. R. 2000. Psychological models of emotion. *The Neuropsychology of Emotion* **137**(3), 137–162.

Scholte, R. H. & De Bruyn, E. E. 2004. Comparison of the giant three and the big five in early adolescents. *Personality and Individual Differences* **36**(6), 1353–1371.

Schwarz, N. & Clore, G. L. 1983. Mood, misattribution, and judgments of well-being: informative and directive functions of affective states. *Journal of Personality and Social Psychology* **45**(3), 513.

Serenko, A., Bontis, N. & Detlor, B. 2007. End-user adoption of animated interface agentsin everyday work applications. *Behaviour & Information Technology* **26**(2), 119–132.

Sherer, K. R. 1999. Appraisal theory. In *Handbook of Cognition and Emotion*, Dalgleish, T. & Power, M. (eds). John Wiley & Sons Ltd, 637–663.

Sloman, A. 1998. Damasio, Descartes, alarms and meta-management. *In SMC'98 Conference Proceedings. 1998 IEEE International Conference on Systems, Man, and Cybernetics*, **3** 2652–2657.

Sloman, A. 2000. Architectural requirements for human-like agents both natural and artificial (what sorts of machines can love?). In *Human Cognition And Social Agent Technology, Advances in Consciousness Research*, Dautenhahn, K. (ed). John Benjamins, 163–195.

Sloman, A. 2002. How many separately evolved emotional beasties live within us? *Emotions in Humans and Artifacts* 35–114.

Smith, C. A. & Lazarus, R. S. 1991. *Emotion and Adaptation*. Oxford Univertity.
Soleimani, A. & Kobti, Z. 2012. A mood driven computational model for gross emotion regulation process paradigm. In *Proceedings of the World Congress on Engineering*, 2.
Steunebrink, B. R., Dastani, M. & Meyer, J. J. Ch. 2012. A formal model of emotion triggers: an approach for BDI Agents. *Synthese* **185**(1), 83–129.
Van Straalen, B., Heylen, D., Theune, M. & Nijholt, A. 2009. Enhancing embodied conversational agents with social and emotional capabilities. In *Agents for Games and Simulations*, Dignum, F., Bradshaw, J., Silverman, B.G. & van Doesburg, W. (eds). Springer Berlin Heidelberg, 95–106.
Thompson, R. A. 1994. Emotion regulation: a theme in search of definition. *Monographs of the Society for Research in Child Development* **59**(2–3), 25–52.
Velásquez, J. D. & Maes, P. 1997. Cathexis: a computational model of emotions. In *Proceedings of the First International Conference on Autonomous Agents*, 518–519. ACM Press.
De Vignemont, F. & Singer, T. 2006. The empathic brain: how, when and why? *Trends in Cognitive Sciences* **10**(10), 435–441.
Watson, D. & Tellegen, A. 1985. Toward a consensual structure of mood. *Psychological Bulletin* **98**(2), 219.
Weiss, G. (ed) 1999. *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. MIT Press.
Wiggins, J. S. 2003. *Paradigms of Personality Assessment*. Guilford Press.
Winikoff, M. 2005. JackTM intelligent agents: an industrial strength platform. In *Multi-Agent Programming*, Bordini, R., Dastani. M., Dix, J. & El Fallah Seghrouchni, A. (eds). Springer US, 175–193.
Yu, C. W. & Choi, J. Y. 2005. Behavior decision model based on emotion and dynamic personality. In *Proceedings of ICCAS'05, Kintex, Gyeonggi-Do, Korea*.
Zadra, J. R. & Clore, G. L. 2011. Emotion and perception: the role of affective information. *Wiley Interdisciplinary Reviews. Cognitive Science* **2**(6), 676–185.
Zajonc, R. B., Murphy, S. T. & Inglehart, M. 1989. Feeling and facial efference: implications of the vascular theory of emotion. *Psychological Review* **96**(3), 395.