Emotional Belief-Desire-Intention Agent Model: Previous Work and Proposed Architecture

Mihaela-Alexandra Puică
Computer Science and Engineering Department
Faculty of Automatic Control and Computers
University Politehnica of Bucharest, Romania

Adina-Magda Florea
Computer Science and Engineering Department
Faculty of Automatic Control and Computers
University Politehnica of Bucharest, Romania

Abstract— **Research in affective computing shows that agents cannot be truly intelligent, nor believable or realistic without emotions. In this paper, we present a model of emotional agents that is based on a BDI architecture. We show how we can integrate emotions, resources and personality features into an artificial intelligent agent so as to obtain a human-like behavior of this agent. We place our work in the general context of existing research in emotional agents, with emphasis on BDI emotional models.**

Keywords— Affective Computing; Agent architecture; Belief-Desire-Intention model

I. INTRODUCTION

When building rational agents, the Belief-Desire-Intention (BDI) model [1] has proven to be one of the best options one can select. It is based on the human reasoning pattern, known as practical reasoning [2]: first decide what one wants to achieve (deliberation), and then decide on how to do it (means-end reasoning). Thus, the agent that follows this model aims at displaying rational behavior by pursuing its *goals*, achieving its *intentions* and using the *beliefs* it has about itself and about the environment.

But research in psychology, neurology and cognitive science shows that people not only use their cognitive functions, but also account for their emotions (even unconsciously) when taking decisions. If these two parts don't interconnect in a proper manner, multiple options are harder to be filtered and bad decisions are easier to take, as stated in work done by Damasio [3]. These results lead to the concept of emotional agents, which aim at being more realistic and providing a more engaging experience in human-computer interaction (HCI), but also at improving the performance of rational agents.

Combining the two concepts, we obtain an agent that reasons based on its beliefs, desires, intentions and emotions. Our aim is to build such an agent, based on previous research done in the field. Consequently, the paper is organized as follows: Section II presents the emotional agent architecture that we developed, and then Section III describes related work concerning existing emotional BDI agents. In Section IV an example of a scenario is proposed and, finally, Section V draws conclusions and outlines future work.

II. EMOTIONAL AGENT ARCHITECTURE

We propose an emotional agent architecture that focuses on the influence of emotions on the behavior and the way resources are used depending on the emotional state. Therefore, we aim to integrate the following *concepts*:

PERCEPTS

- anything that comes from the environment: stimuli or messages from other agents
- influenced by emotions

BELIEFS

- acquired from percepts
- revised to account for current beliefs and new percepts
- influenced by emotions

DESIRES

- goals received by the agent at design time
- constant over time

OPTIONS

- alternatives to accomplish the desires
- generated based on current beliefs and intentions
- influenced by emotions

INTENTIONS

- options that the agent has committed to
- revised based on current intentions, beliefs, options, emotions and available resources
- open-minded commitment - agent is committed to the intention as long as it is not achieved yet, it is not believed impossible to achieve and it is still a goal for the agent

EMOTIONS

- primary emotions may determine instinctual behavior
- secondary emotions influence cognitive processes and available resources
- fixed set of emotions for each scenario

PERSONALITY

- two axes: extrovert-introvert, psychologically stable-unstable

- four types: sanguine, choleric, melancholic, phlegmatic

RESOURCES

- maintained in a structure that gives access to them selectively, based on emotions

- fixed set of resources for each scenario

The *functions* that define the processes that take place within the agent internal mechanisms are the following:

- $perceive: Env \; x \; E \to P$ – perceives the stimuli in the environment (Env), influenced by emotions (E); returns the *percepts P*

- $peu: P \; x \; E \; x \; I \to E$ – primary emotions update; the new percepts, influenced by intentions and current emotions, generate *primary emotions*

- $react : E \; x \; P \; x \; AR \to \pi$ – defines reactive behavior; reacts to the percepts, influenced by emotions and guided by the *available resources* (AR), returning a *plan* to be executed (generally consisting in only one simple action)

- $brf: P \; x \; B \; x \; E \to B$ – *belief* revision function; revises current beliefs based on percepts and influenced by emotions

- $seu: B \; x \; E \; x \; I \; x \; AR \to E$ – secondary emotions update; beliefs, intentions, available resources and current emotions generate *secondary emotions*

- $ru: R \; x \; E \to AR$ – resources update; emotions influence the *available resources* the agent can use from the set of all *resources* it has (which are not always accessible)

- $analyzer: B \; x \; D \; x \; I \; x \; E \to O$ – appraises the current situation in the context of its beliefs, desires and intentions, influenced by its emotions, to generate the current *options* that it has to cope with the situation

- $filter: O \; x \; B \; x \; E \to I$ – filters the available options to find the *intentions* that it will be committed to

- $plan: I \; x \; AR \to \pi$ – structures intentions into plans, according to the available resources

- $execute: \pi \to Env$ – executes the plan

- $reconsider: I \; x \; B \to \{true, false\}$ – decides whether it is necessary to reconsider or not the intentions

- $sound: \pi \; x \; I \; x \; B \to \{true, false\}$ – decides whether executing the plan will lead to intentions achievement or not

Fig. 1 illustrates the proposed architecture. The processes that take place within the agent control loop are shown in Fig. 2. In what follows, we will explain in detail this process,

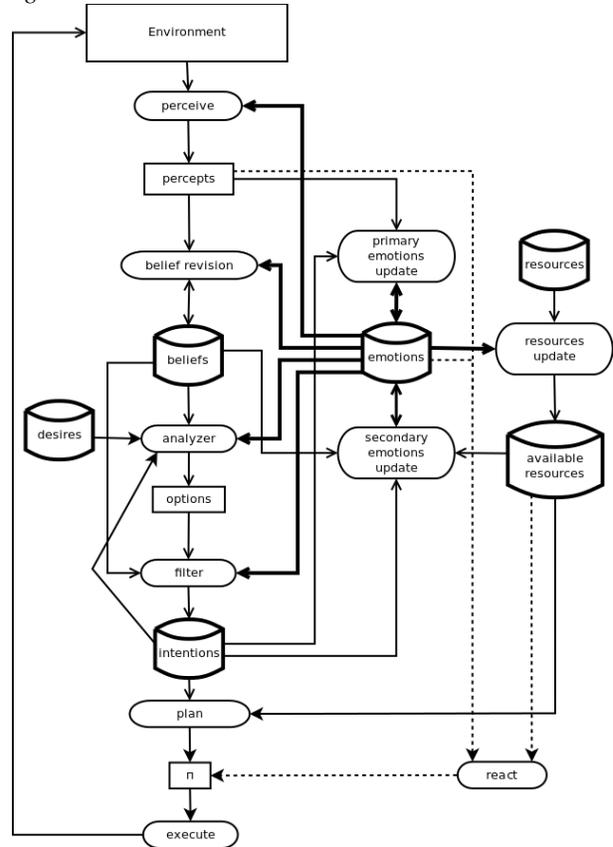*walking step by step through the internal mechanism of the agent.*

Fig. 1.   Proposed Architecture

First, the agent perceives the stimuli in the environment ($perceive$), which can be the occurrence of an action, a change in the state of the world or a message sent by another agent. Percepts are then quickly appraised at an affective level and may determine the experience of primary emotions (peu).

When the intensity of the new emotions passes a certain threshold ε, primary emotions may determine the agent to execute some predefined reactive behavior. For example, strong fear can determine a sudden fall back or a sudden stroke on what caused the fear (depending on the personality). Percepts give the context and available resources determine the action that is to be taken. The $react$ function returns a plan π which generally contains only one action.

Primary emotions generate a reactive behavior which could be unnecessary, but it is very useful in survival situations. Fig. 4 shows a dotted input for the react function to emphasize that the process is not guaranteed to happen at every step, but it is a matter of strong, sudden emotions.

After the initial instinctual behavior, the percepts influence belief revision (br). This process must account for both current beliefs and new percepts, acting like a truth maintenance system which outputs consistent beliefs:

After belief revision, secondary emotions are being computed. These emotions are the result of a cognitive

appraisal of the situation which takes into consideration not the raw percepts, but the newly revised beliefs.

```
1.  D = D₀;
2.  while true do
3.      P ← perceive(Env, E);
4.      E ← peu(P, E, I);
5.      if intensity(E) > ε then
6.          π ← react(E, P, AR);
7.          P ← perceive(Env, E);
8.      end if
9.      B ← brf(P, B, E);
10.     E ← seu(B, E, I, AR);
11.     AR ← ru(R, E);
12.     O ← analyzer(B, D, I, E);
13.     I ← filter(O, B, E);
14.     π ← plan(I, AR);
15.     while not (empty(π)
                  || succeeded(I, B)
                  || impossible(I, B)) do
16.         α := head(π);
17.         execute(α);
18.         π := tail(π);
19.         P ← perceive(Env, E);
20.         E ← peu(P, E, I);
21.         if intensity(E) > ε then
22.             π ← react(E, P, AR);
23.             P ← perceive(Env, E);
24.         end if
25.         B ← brf(P, B, E);
26.         E ← seu(B, E, I, AR);
27.         AR ← ru(R, E);
28.         if reconsider(I, B) then
29.             O ← analyzer(B, D, I, E);
30.             I ← filter(O, B, E);
31.         end if
32.         if not sound(π, I, B) then
33.             π ← plan(I, AR);
34.         end if
35.     end while
36. end while
```

Fig. 2.   Agent Control Loop

But emotions in their turn influence the available resources that an agent perceives (ru), a process that models human resource usage: when experiencing fear, one can run faster or be stronger. Thus, at a given moment, depending on the affective state, an agent can have access to different resources.

Next follows deliberation (analyzer and filter) and means-end reasoning (plan). The *analyzer* function determines the available options that the agent has to accomplish its desires. The *filter* function chooses from the available options those that the agent will be committed to. The selected options are the agent intentions.

After deciding on the intentions, the agent must *plan* on how to achieve them, taking into consideration available resources. In the end, the plan is ready to be executed (*execute*).

But during this process, the agent still has to pay attention to the stimuli in the environment. If a plan contains several actions, the environment can change its state during execution, so the agent should not omit to reconsider its plan or its intentions. In the light of new events, the intentions could no longer be necessary, so the agent should consider dropping them. This is done in the *reconsider* function.

If this functions returns true, then the agent restarts the deliberation process, returning renewed intentions. This approach is known as open-minded commitment strategy relative to intentions. In addition, the agent must test the soundness of the plan in relation to the intentions, and replan if necessary (*sound*).

III.   MODELING EMOTIONS

A. Emotions

We define emotions using the circumplex model of affect developed by James Russell in 1980 [13]. This is a dimensional model that represents affect in a 2D emotional space determined by the valence and arousal axes. Fig. 3 shows eight emotions placed in this space. The valence axis is defined by pleasure-misery, while the arousal axis by arousal-sleepiness. The other four emotions simply define the contents of each quadrant.
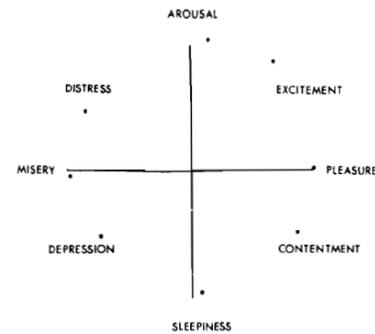

Fig. 3.   James Russell's Circumplex Model Of Affect

Consequently, we thus can assume that an emotion is a point in this 2D space given by the pair (v, a), where both v and a are rational numbers within given intervals: v is the value of the valence and a is the value of the arousal. Following the OCC model of emotions [12], an event is appraised in terms of beliefs, desires and intentions, returning a certain score regarding its valence (positive or negative) and arousal (the intensity of emotion felt).

We believe that the circumplex model of affect permits to easily define different sets of emotions for different scenarios. The six basic emotions are not suitable in every context; for example, in a fire scenario it is less likely that one will feel joy. Most probably, a drama scenario may activate emotions in the upper left quadrant, while a romance scenario may activate emotions in the lower right quadrant.

B. Resources

Resources have a central role in decision making because decisions are based on the available resources that the agent knows it has. In a critical situation, people may suddenly appear to have more energy than they usually do. The arousal opens new paths in the person's capabilities, but working at full capacity exhausts the energy, so it is not a desired feature unless situation requires it.

In our architecture, we define three types of resources, each with different access conditions. Type A of resources can always be used by the agent. Type B defines resources that the agent can access only when its emotional intensity goes over a given threshold. Type C of resources is used only in survival situations. Fig. 4 shows the relation between arousal and resource accessibility, using the Yerkes-Dodson law [14] to emphasize the fact that using more resources may improve performance.

The picture depicts the emotion intensity thresholds which determine the type of resources the agent has access to. Thus, if we define by e_i the emotion intensity, and by p the performance, we have the following cases:

$$0 < e_i < e_{i1} \rightarrow \begin{cases} access\ to\ type\ A \\ p \le p_1 \end{cases}$$

$$e_{i1} < e_i < e_{i2} \rightarrow \begin{cases} access\ to\ type\ A\ \&\ B \\ p_1 < p \le p_2 \end{cases}$$

$$e_{i2} < e_i < e_{i3} \rightarrow \begin{cases} access\ to\ type\ A\ \&\ B\ \&\ C \\ p_2 < p \le p_{max} \end{cases}$$

$$e_{i3} < e_i < e_{i4} \rightarrow \begin{cases} access\ to\ type\ A\ \&\ B \\ p_1 < p \le p_2 \end{cases}$$

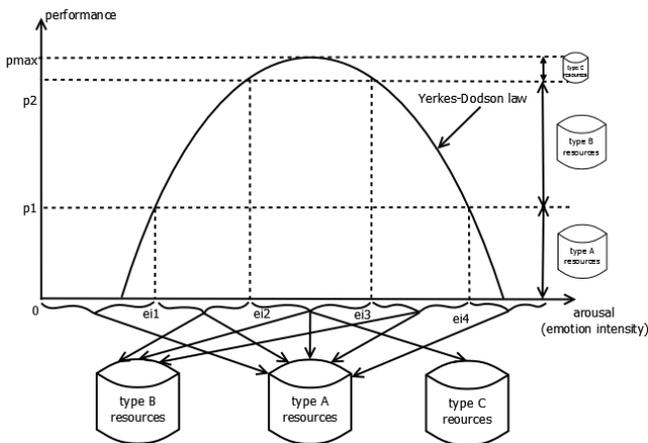$$e_i < e_{i4} \rightarrow \begin{cases} access\ to\ type\ A \\ p \le p_1 \end{cases}$$

Fig. 4.   Agent Accessibility To Its Resources

C. Personality

We model personality through the four temperament types: sanguine, choleric, melancholic and phlegmatic. These are represented by the four areas delimited by two axes in a 2D space: extraversion axis (extrovert/introvert) and neuroticism axis (psychologically stable/unstable). We chose this model because of its simplicity and of a correspondence that can be made between the two axes and the parameters in our agent architecture. The mapping is shown in Fig. 5 and is explained below.
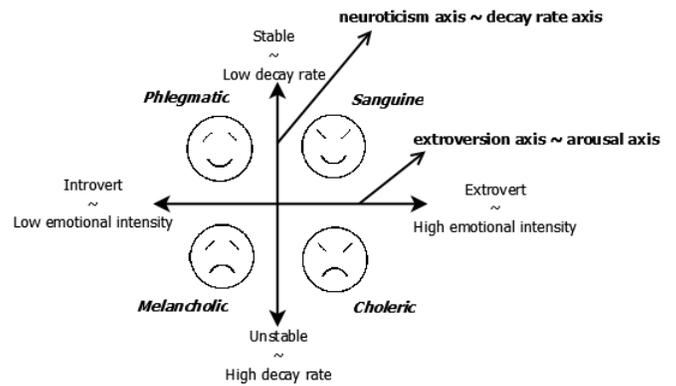
Fig. 5.   The four temperaments

The introvert/extrovert axis is associated with the arousal (or emotional intensity). According to Eysenck [15], introverts are characterized by higher cortical arousal than extroverts, therefore the former require less stimuli from the environment (less emotional intentisy), while the latter seek situations which stimulate them more (give higher level of emotional intensity). The psychologically stable/unstable axis is associated with the decay rate of agents' emotions. Thus, a stable agent has a low decay rate, with a coherent affective state; an unstable agent has a high decay rate, quickly moving from one emotional state to another. Resuming, the four temperament types have the characteristics shown in TABLE 1:

TABLE 1. AGENT PERSONALITY

Temperament	Extraversion	Neuroticism	Characteristics
Sanguine	extrovert	stable	-high emotion intensity -low decay rate
Choleric	extrovert	unstable	-high emotion intensity -high decay rate
Melancholic	introvert	unstable	- low emotion intensity - high decay rate
Phlegmatic	introvert	stable	-low emotion intensity -low decay rate

The role of personality in the architecture is to define how emotions are updated. This is done through two variables: emotion intensity and decay rate. First of all, each agent has a predefined value for arousal: the sanguine and choleric have a lower value, the melancholic and phlegmatic have a higher value. When an event takes place or the state of the environment changes, each of this adds a certain emotional intensity to the default value of arousal. Thus, in the 2D emotional space the emotion point is shifted up for the introverts over the extroverts. For example, if the execution of the plan fails, the extrovert might be annoyed, but the introvert might be angry (annoyed and angry represent the same basic emotion, but at different intensities). The second variable is also predefined and represents the value that the emotion is decayed with at each step. If an agent is angry, but then an intention is achieved, the melancholic and the choleric might forget the anger, feeling joyful, but the sanguine and the phlegmatic will not forget the anger so rapidly and will not pass so rapidly to the joyful state.

Personality also influences the performance of the agent. To capture this, we need to take a look at Yerkes-Dodson Law of arousal [14]. This law states that performance increases with the increase of arousal, but up to a point where the performance is maximum. After that, performance drops if arousal keeps increasing. The value of arousal for which performance is optimum is different for different personalities. Thus, for extroverts the graph is shifted to the right comparing to the graph for introverts, as shown in Fig. 6.
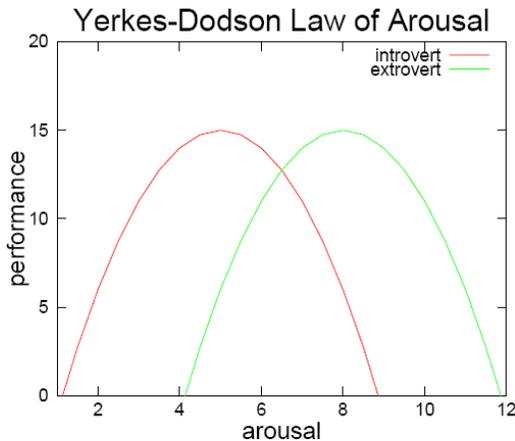

Fig. 6.   The Yerkes-Dodson Law Of Arousal For Introverts And For Extroverts

IV.   PREVIOUS WORK

Early work was done by Padgham and Taylor [5]. They built a system in which the reasoning is done by the Distributed Multi Agent System (dMars) [4]. What is specific for their model is that each agent has a given threshold for the emotional gauge, and when this is passed, the *belief* that the agent feels that particular emotion *is asserted*. Therefore, the main problem of this architecture is that emotions do not directly influence the reasoning process, but are considered only as beliefs. Nevertheless, the first step of integrating agent emotions and BDI agent model into the same system was done. Consequent research goes into more detail on the relationship between emotions and the reasoning process.

Pereira et al. [6] describes a conceptual model of an Emotional BDI agent which aims at identifying the disadvantages of the BDI model and to overcome them. The solution was to add new concepts to the model. Thus, their architecture contains resources, capabilities, a module for sensing and perception and a module for managing the emotional state. While resources and capabilities are simply "internal representations of the means that an agent has to execute upon its environment" (concrete means and abstract plans of action, respectively), the emotional state manager is a module that generates the emotions felt by the agent. Nevertheless, it is not fully specified, the authors mentioning only the main characteristics that this module should include: a well-defined set of Artificial Emotions, various triggering events and a decay rate for each Artificial Emotion. Although the details of the emotional module are left for further consideration, the idea of introducing resources and capabilities in the BDI model is worth to be considered.

The following year, independent of [6], Parunak et al. [7] describes the DETT model for situated agents applicable in the particular situation of a combat. Their aim is to simulate faster than real-time a large number of combatants. The architecture includes two reasoning processes (appraisal and analysis) and four new concepts (Dispositions, Emotions, Triggers and Tendencies). The appraisal process assesses the beliefs in the context of agent *disposition* and returns the *emotion* felt. The 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. In their turn, emotions impose a *tendency* on the intentions that are to be selected by the analysis process. Although it is important for an agent to be able to act in real time, we would like to build an architecture that is not specific to a particular situation.

A more sophisticated model was presented by Jiang et al. [8]. They focus on the influence of emotions on the decision making process. The novelty of their EBDI architecture consists in considering primary and secondary emotions, as well as three sources for beliefs: perception, communication and contemplation. Primary emotions are connected to reactive behavior and fast decisions. If time permits, further deliberation is conducted and secondary emotions generated. The authors believe resources and capabilities introduced by Pereira et al. [6] to be unnecessary. Instead, they solve the resource boundary condition by adding priorities for beliefs, desires, intentions and emotions, deleting those with lowest priority when memory is full. On the other hand, they believe the problem of reconsideration to be context specific, so they leave it for the plan execution function to deal with it. The authors also criticize [6] for not emphasizing the differences between the emotional agent and other agents. In contrast, they have implemented their architecture and tested it in Tileworld using one emotional agent and two non-emotional agents. Results show that the first has better results than the latter.

The most recent emotional BDI architecture is that of Jones et al. [9], called PEP-BDI. The authors criticize Jiang et al. [8] for missing personality and physiology aspects and Parunak et al. [7] for modeling only two emotions in relation with two personality aspects. Their main motivation lies in the context of global security, and their goal is to simulate a crisis situation in a virtual reality environment. They add emotions, personality and physiology to the BDI model, stating that these are key concepts in the context of crisis management. The paper illustrates the decision making process on a scenario which involves escaping from a fire. One should note that physiological parameters are not necessary in a non-crisis scenario, but resources are a more suitable component to consider.

In parallel with the above mentioned systems, two more architectures were developed, both in the topic of emotional BDI models, but each one independent of previous works and each one independent of one another.

One work is that of Hernandez et al. [10]. The authors develop a modular architecture, named BDIE, which contains four separate modules: Perceptual System, Motivational System, Behavior System and Emotional System – mapped to Beliefs, Desires, Intentions and Emotions, respectively.

Beliefs, evaluators and emotions are organized on two levels. First level beliefs are acquired through perception and are then affectively appraised by the first level evaluators, giving the primary emotions (fear, surprise). If one of these is active, control passes to the planning algorithm (contained in the Behavior System). Otherwise, second level beliefs are inferred from the first level beliefs and are cognitively appraised by the second level evaluators, giving the secondary emotions (happiness, sadness, anger). Desires comprise goals (need to be achieved) and homeostatic variables (need to be maintained). However, the paper doesn't explain how the agent chooses its desires and intentions; in the scenario considered the only desire is survival and the only behavior is to display facial expressions according to the emotions induced by certain pictures. The modular structure of the system can be appointed as the most interesting aspect of the BDIE architecture.

The second architecture, called BDE, was developed by Florea and Kalisz [11] on the hypothesis that emotions contribute to behavior anticipation, which in turn contributes to making an agent more realistic. The authors show that an event may generate several emotions, which are then integrated into an emotional state using a rule-based approach. Also, the authors account for both emotion decaying and emotional memory; thus, powerful emotions remain in memory even if they are momentarily decayed. The paper specifies the emotion eliciting conditions (EEC) and the influence of emotional state on behavior (IEB) for seven emotions. Although the architecture is not very complex, by not considering all the details of the human mechanisms, the simplicity of the BDE model makes it more appropriate in not so dynamic environments.

TABLE 2 summarizes the characteristics of the presented models. In what follows, we detail the similarities and differences among them.

First of all, one can note that each paper is concerned with different issues:

- Pereira et al. [6] - BDI improvement

- Parunak et al. [7] - faster-than-real-time, large number of combatants

- Jiang et al. [8] - improved decision-making

- Jones et al. [9] - handling crisis situations

- Hernandez et al. [10] - performance, HCI

- Florea and Kalisz [11] - behavior anticipation

Secondly, there are also obvious differences with respect to the architecture components. In [8] beliefs are acquired using three different methods: stimuli in the environment, messages from other agents and contemplation. [6], [7], [10] and [11] don't take into consideration communication, but use one belief revision function that includes both new percepts and current beliefs. The same goes with [9], but percepts can be obtained through stimuli, messages and physiological parameters (which define the agent's health).

Next, the reasoning process happens differently. In [7] desires are constant over time, while in [11] emotions (and only emotions) influence them. The deliberation process updates the intentions, and the means-end reasoning process structures these intentions into plans. For [11] and [7], desires are predefined and need to be filtered only if they are inconsistent.

A set of consistent desires forms the goals of the agent. [6], [8] and [9] consider desires as options that need to be constantly generated based on current desires and beliefs. In [8] and [9], deliberation process is formed by the options function (generating desires) and the filter function (generating intentions); the means-end reasoning is done by the plan function, which structures intentions into plans. In [6] on the other hand, the first process that occurs is generate options, which maps to means-end reasoning; this generates desires and intentions that hierarchically flow from abstract to concrete, until executable actions are obtained. Then, the deliberation process (filter) chooses those intentions that the agent will be committed to.

What all the papers have in common is the focus on the influence of emotions over deliberation and means-end reasoning. How the affective state affects the planning algorithm is not a priority for any of the authors. The commitment strategy is yet less discussed, although the trade-off between reconsideration and the degree of commitment is an important aspect of practical reasoning. From the agent control loops presented it can be inferred that [11] uses an open-minded strategy (the agent is committed to an intention as long as it is still believed possible), while [6], [8] and [9] use a single-minded strategy (the agent is committed to an intention until it is either realized or not possible).

Last, but not least, [6] only describes a conceptual architecture, [9] shows a fire scenario and [11] exemplifies the emotion generation and influence over behavior for seven basic emotions. On the other hand, [7], [8] and [10] implement their agents and show experimental results. Thus, [8] runs both emotional and non-emotional agents in Tileworld, showing that the EBDI agent has a better performance; [10] implement their architecture on a robotic head which changes its facial expression based on color and luminance of an image; and [7] creates an architecture which is specific to a combat situation, mapping to concrete beliefs, desires and emotions.

We built our architecture based on the work presented above, focusing on agent performance, on simulating human mechanisms of internal resource usage and on agent personality. We were mostly attracted by the EBDI model [8] for emphasizing the influence of emotions on decision making and for taking into account both primary and secondary emotions. Additionally, we are interested in the model of Pereira et al. [6] for adding resources and in the PEP→BDI model [9] for considering the personality influences on agent behavior. Ultimately, we valued the usage of the OCC theory on beliefs, desires and intentions in the BDE model [11].

TABLE 2. COMPARISON OF EXISTING EMOTIONAL BDI SYSTEMS

Components / Model	Belief	Desire	Intention	Emotion	Other
Pereira et al., 2005 [6]	- acquired through perception &inference - one revision function	- revised through distinct algorithms, depending on resources, capabilities and emotional state	- revised through distinct algorithms, depending on resources, capabilities and emotional state	- Emotional State Manager - not detailed	- resources - capabilities - Sensing and Perception Module
Parunak et al., 2006 [7] (DETT)	- mapped to digital pheromones	- wants - constant over time	- result of analysis process - tendency imposed by emotions	- result of appraisal process (OCC) - triggered by beliefs, depending on disposition	- Disposition, Emotion, Trigger, Tendency
Jiang et al., 2007 [8] (EBDI)	- acquired through perception, communication & contemplation - three revision functions	- not influenced by emotions	- influenced by emotions, desires & intentions - state of affairs that the agent has committed to achieve	- primary - secondary	-
Jones et al., 2009 [9] (PEP-BDI)	- acquired through perception, communication &physiology - one revision function, three perception functions	- options - influenced by beliefs, intentions, personality and physiology, but not by emotions	- filtered options - influenced by beliefs, desires, intentions, emotions and physiology, but not by personality	- primary - secondary	- Personality, Emotion, Physiology
Hernandez et al., 2004 [10] (BDIE)	- first level (acquired through perception) - second level (acquired through first level belief revision)	- goals (to be achieved) - homeostatic variables (to be maintained)	- contains the planning algorithm - connected with emotions through goals	- primary & secondary - emotional space divided into emotional sectors	-
Florea and Kalisz, 2004 [11] (BDE)	- acquired through perception & inference - one revision function	- influenced only by the emotional state	- course of action to be taken in order to achieve desires	- OCC - emotional state - emotional memory - decay	-

V. SCENARIO

The story is that of a man that wants to gather the garbage in a city, with garbage being well-defined items that can be collected by that man. Because he is walking, he risks to be bitten by dogs that have no master and wander freely on the streets. He thus has to be careful to avoid the dogs. The man doesn't know where the garbage items are, new items being possible to appear anytime and anywhere in the city, but other people may tell him about certain item locations.

We can model this story with an agent situated in a dynamic environment, a 2D map which contains artifacts, but also traps. The agent has the goal of gathering as many artifacts as possible, taking care to avoid the traps. The dynamism of the environment is given by the random appearance of the artifacts and by the random appearance, disappearance and moving of the traps. The agent can see only a fixed number of cells in front of him, but it can remember the position of the artifacts and the traps.
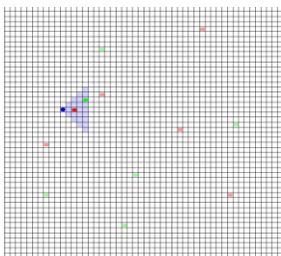

Fig. 7. Example Of A Map For The Proposed Scenario

In Fig. 9, the agent is represented by the blue circle. The light blue triangle besides the agent shows what the agent can see from its current position. The green ovals are the artifacts and the red ovals are the traps. The artifacts appear random on the map. If their number decreases below a specified value (because of being collected by the agent), other artifacts randomly appear on the map. Traps may appear and disappear randomly, and they also have a random move (they are more like non-rational agents which wander on the map with no purpose whatsoever). An agent is said to *fall into a trap* if it steps into a cell simultaneously with a trap. If this is the case, its energy level decreases. Traps cannot appear on the cell that the agent is currently standing, neither on a cell where there is already an artifact. TABLE 3 shows how the components integrated into the architecture are mapped to the current scenario.

In this scenario, resources refer to the possible actions that an agent may take. The agent has a specified energy level that is decreased when using resources (more or less and each type of resource used consumes more or less of this energy. The same energy is consumed when falling into a trap. The agent has the possibility to restore its energy level by taking the action *sit* (analogous to the situation in which one has to relax to recover from a great effort).

Depending on the personality, the agent may pass from one emotion to another faster or slower. For example, if it fell into a trap and is angry, but shortly after it collects an artifact, a choleric agent will be joyful, forgetting the anger, while a phlegmatic one will feel only slight joy, still alert to traps (low

decay rate for phlegmatic, high decay rate for choleric). The belief that a trap is close will generate fear in a melancholic agent, but a sanguine agent will feel fear only if the trap is much closer (the melancholic has a high default arousal, so small increase in emotion intensity will make it alert, while the sanguine needs higher arousal level to actually feel the new emotion).

TABLE 3. INSTANTIATION OF THE AGENT COMPONENTS

Percepts	What the agent has in sight (clear cell, artifact, trap) Messages from other agents containing artifacts or traps coordinates (not shown in the figure)
Beliefs	Agent coordinates Artifact coordinates Traps coordinates
Desires	Collect artifacts Avoid traps
Options	Move to the artifacts with known position
Intentions	Move to the selected artifact(s)
Emotions	Joy – artifact collected (goal accomplished) Fear – trap near (negative belief) Anger – fall into trap (goal not accomplished) Relief – trap disappears from sight (positive belief)
Personality	One of Choleric / Melancholic / Sanguine / Phlegmatic
Resources	Possible moves: Type A: turn 90°, move forward 1 cell, sit Type B: jump left / right / behind 1 cell Type C: jump forward over trap (2 cells)

VI. CONCLUSION

In this paper we presented an emotional BDI architecture which accounts for agent resources and personality. We presented the agent control loop, the emotional mechanisms, the resource usage and the personality influence on agent behavior. The main contribution is that the agent built using this architecture has both cognitive and reactive behavior. The BDI architecture focuses on the cognitive functions of the brain, but in a very dynamic environment the reactive behavior gives the agent the possibility to quickly deal with unexpected events. For this reason, we believe our architecture should improve agent performance. Moreover, by considering the influence of both primary and secondary emotions, we can faithfully replicate the human reasoning process, which is a step forward into giving agents human intelligence.

We based our work on previous research, which we have reviewed briefly, showing similarities and differences and noting the features that inspired us. Thus, we were specifically attracted by the EBDI model [8], because of its proved improved performance; also, the model of Pereira et al. [6] is interesting for including resources in the architecture and the PEP→BDI model [9] for including personality; last, but not least, the cognitive appraisal theory applied on beliefs, desires and intentions, as described in the BDE model [11].

Further on we need to implement this architecture and to actually test it on the described scenario. Other scenarios will also have to be developed and implemented, so that the agent can be tested in different situations.

REFERENCES

[1] Rao, A. and Georgeff, M. (1991). Modeling rational agents within a BDI-architecture. In Proceedings of Knowledge Representation and Reasoning(KR&R-91), pages 473–484.

[2] Bratman, M., Israel, F., and Pollack, M. (1988). Plans and resource-bounded practical reasoning. Computational Intelligence.

[3] Damasio, A. (1994). Descartes' Error: Emotion, Reason and the Human Brain. Gosset/Putnam Press.

[4] D'Inverno, M., Luck, M., Georgeff, M., Kinny, D. and 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.

[5] Padgham, L. and Taylor, G. (1997). A system for modelling agents having emotion and personality. In Proceedings of PRICAI Workshop on Intelligent Agent Systems: Theoretical and Practical Issues, Lecture Notes in AI 1209, pages 59–71. Springer-Verlang.

[6] Pereira, D., Oliveira, E., Moreira, N., and Sarmento, L. (2005). Towards an architecture for emotional BDI agents. In IEEE Proceedings of the 12th Portuguese Conference on Artificial Intelligence (EPIA'05), pages 40–47. Springer.

[7] Parunak, H. V. D., Bisson, R., Brueckner, S., Matthews, R., and Sauter, J. (2006). A model of emotions for situated agents. In Proceedings of 5th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'06), pages 993–995.

[8] Jiang, H., Vidal, J., and Huhns, M. (2007). EBDI: An architecture for emotional agents. In Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'07), pages 38–40.

[9] Jones, H., Saunier, J., and Lourdeaux, D. (2009). Personality, emotions and physiology in a bdi agent architecture: the PEP→BDI model. In Proceedings of WI-IAT - IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology, volume 2, pages 263–266.

[10] Hernandez, D., Deniz, O., Lorenzo, J., and Hernandez, M. (2004). Bdie: a bdi like architecture with emotional capabilities. In Papers from AAAI Spring Symposium: Architectures for Modeling Emotion: Cross-Disciplinary Foundations.

[11] Florea, A. M. and Kalisz, E. (2004). Behavior anticipation based on beliefs, desires and emotions. International Journal of Computing Anticipatory Systems, 14:37–47.

[12] Ortony, A., Clore, G., and Collins, A. (1988). The cognitive structure of emotions. Cambridge University Press.

[13] Russel, J. (1980). A circumplex model of affect. Journal of Personality and Social Psychology, 39(6):1161–1178.

[14] Yerkes, R. and Dodson, J. (1908). The relationship of strength of stimulus to rapidity of habit formation. Journal of Comparative Neurology and Psychology, (18):459–482.

[15] Eysenck H. (1967). The biological basis of personality. Springfield, IL: Thomas Publishing.

AUTHORS PROFILE

Mihaela-Alexandra Puică received her BSc in Computer Science in 2009 and MSc in Artificial Intelligence in 2011, both from University Politehnica of Bucharest, Faculty of Automatic Control and Computers. She is now a PhD student at the same university, doing her research in the field of Affective Computing, both emotion simulation and emotion recognition. She is also a member of the Artificial Intelligence and Multi-Agent Systems (AI-MAS) Laboratory.

Adina-Magda Florea is a Professor at the Computer Science and Engineering Department from the University Politehnica of Bucharest, Faculty of Automatic Control and Computers. She is teaching Artificial Intelligence, Machine Learning, Knowledge Representation and Reasoning and Multi-Agent Systems. She is head of the Artificial Intelligence and Multi-Agent Systems (AI-MAS) Laboratory. Her research interests include Multi-Agent Systems, Ambient Intelligence, Affective Computing and E-learning.