# Behavior Anticipation Based on Beliefs, Desires and Emotions

Adina Magda Florea, Eugenia Kalisz
Department of Computer Science
"Politehnica" University of Bucharest
Spl. Independentei 313, Bucharest, Romania
E-mail: adina@cs.pub.ro, ekalisz@cs.pub.ro
http://turing.cs.pub.ro/ai_mas

**Abstract**
Most of the existing models of intelligent software agents fail to consider an important aspect of human behavior, namely the impact of emotions on processes such as motivation, decision-making, planning, learning, and anticipation. The paper presents an emotional reasoning model of artificial agents, called Belief-Desire-Emotion (BDE). The model is built upon the influential Belief-Desire-Intention agent architecture and follows the Ortony-Clore-Collin's cognitive appraisal theory of human beings. We describe the different stages of the emotion generation process and emphasize how this process influence theoretical reasoning, such as belief and desire revision, and practical reasoning, such as means-end analysis. Additionally, we propose a set of basic emotions, and we exemplify how they are generated and the way they influence the behavior of our BDE agents.

**Keywords**: agent-based systems, model of anticipatory and predictive capabilities, BDI model, emotions, emotional intelligence.

## 1 Introduction

There is a steady growing interest in the topic of intelligent agents, both from the point of view of expanding the possible application areas and from the one of modeling human behavior and intelligent problem solving. Bertil Ekdahl [1] says that in order to be a good human agent it is necessary to be able to guess and foresee the future, that is, to anticipate what will happen and act accordingly. If the analogy with human beings is to be followed, a software agent should be evaluated from the point of view of its anticipatory abilities.

Current models of intelligent (software) agents are limited by their inability to model realistic human behavior by ignoring, most of the time, a set of moderators that influence this behavior: emotions, level of stress, individual differences, etc. Recently, research was undertaken towards adding another human-like dimension to artificial intelligent agents, namely emotions, aiming to achieve both the artificial synthesis of emotions and the recognition of human emotions. Emotions often play an important role in human's behavior, for example in decision-making, motivation, judgment, learning, etc. The thesis of this paper is that emotions can contribute to a more accurate modeling of the anticipated behavior of artificial intelligent agents.

**International Journal of Computing Anticipatory Systems, Volume 14, 2004**
**Edited by D. M. Dubois, CHAOS, Liège, Belgium, ISSN 1373-5411 ISBN 2-930396-00-8**

Starting from the well-known and influential BDI (Belief-Desire-Intentions) model of agency [2], the paper proposes a BDE (Belief-Desire-Emotion) model of agents' behavior by modifying the BDI architecture to account for the influence of emotions and emotional intelligence, based on Ortony's event-appraisal theory. It also proposes a set of basic emotions together with the mechanisms used to appraise events and influence behavior in the context of the BDE model.

## 2 Existing emotional models

Psychological theories, called also "interactionalist" theories, emphasize that emotions arise from the interaction between the external environment and a person's internal disposition. Models of emotions typically use the term *appraisal* to refer to the process of qualitatively evaluating the significance of events. The most prominent model is Andrew Ortony, Gerald Clore, and Allan Collin's **cognitive appraisal theory**, commonly referred to as the **OCC** model [3]. This theory views emotions as arising from a valued reaction to events in the light of the agent's goals, standards, and attitudes. For example, a student having the goal to play tennis would be happy if today's class is cancelled and gives him the opportunity to achieve his goal.

Computer simulated emotions may be classified in two main categories: one in which an artificial agent deliberately convey emotions, such as a virtual human choosing its emotional expression based on its desired impact on the user, and simulation-based approaches aimed to simulate the mechanisms by which emotions arise. The second category of approaches is mainly based on the OCC model. The OZ project [4] simulates believable and emotional social agents, which perceive events in an environment. The event is evaluated according to the agent's goals, standards and attitudes and special rules are used to produce an emotion with a specific intensity. The Affective Reasoner project [5] is also based on the OCC model. The agents are able to produce twenty-four different emotions and can infer other agent's emotional states from the emotional expression and actions of other agents.

The PETEEI project [6] simulates the emotional process as a web-like structure, which interacts with motivation, behavior and learning. Starting also from the OCC model, PETEEI appraises emotions in terms of events' desirability and expectations. The desirability of an event is learned by reinforcement learning and user's patterns of behavior (event expectations) are learned using a probabilistic approach based on frequencies of actions. The presented models are good simulations of emotions in the context of agent's motivations and social interactions but they use a too simplistic approach for the agent modeling and reasoning capabilities. The Emile project [7] proposes a more comprehensive approach to simulate links between emotion simulation and agent's reasoning in the context of military training simulations. It differs from other computational models of emotions by emphasizing the role of plans in emotional reasoning. Emile addresses the issues of how to develop and execute plans to achieve goals, how to model the plans and emotional responses of other synthetic agents, and how to guide the presentation of information through emotional expressions or gestures. Emile agents are definitely more accurate from the point of view of their representation and reasoning capabilities, but they focus only on the planning reasoning abilities of agents, disregarding some other important attributes and forms of reasoning such as beliefs and desires.

Our proposed model of emotional intelligence starts also from the appraisal theories of emotion creation but places the appraisal of an event in the context of the agent's beliefs, desires, and intentions. The agent behavior is based on the BDI model [2], a well-know and quite used model of intelligent agents behavior, and the emotional reasoning is interlinked with both the theoretical reasoning of the agents [8], such as belief revision, and the agent's deliberation process towards satisfying its desires and goals.

## 3 Belief-Desire-Emotion model

The BDI (Belief-Desire-Intention) model [2] is a high-level specification of a practical architecture for a resource-bounded agent. It performs belief revision, weighting of competing alternatives, and means-end analysis. A BDI agent has a set of **beliefs** ($B$) - information the agent has about the world; a set of **desires** ($D$) - the state of affairs that the agent would wish to bring about; and a set of **intentions** ($I$) - options for accomplishing the desires that the agent has committed to achieve. Intentions play a critical role in practical reasoning as they limit the available options [9] and are structured into plans by means of a planner (the means-end reasoner). The agent has a set of percepts by means of which it recognizes an **event** that occurred in the environment, event that can be either a change in the state of the world or the occurrence of an action.

The structure of a BDI agent is depicted in Figure 1, where the three types of agent's reasoning have been denoted by the following functions:

- *brf:* $B \times p \rightarrow B$ is the belief revision function, with $p$ the set of percepts;
- *options:* $B \times D \times I \rightarrow I$ is the function which weights competing alternatives to achieve the desires and decides the course of action to be taken;
- *plan:* $B \times I \rightarrow \Pi$ is the function that structures intentions into plans, with $\Pi$ being the possible plans an agent has, depending on its available actions (competencies).

The BDI model can account for the anticipatory character of an intelligent agent. The desires represent a set of potential future states anticipated by the agent, of which it selects the state(s) it considers achievable and builds the appropriate conceptual model(s), by applying the function options. The function plan builds a concrete model for achieving resulting intentions, using both externalist and internalist anticipation [10]. Considering the desires as potential future states of the agent and the distinction between strong and weak anticipatory systems in [11], a BDI artificial agent can be considered a strong anticipatory system as it determines its current state based on states at past times (intentions), present time (actions in plans), and even future time (unaccomplished desires). Moreover, the decision mechanism is embedded in the agent itself [12].

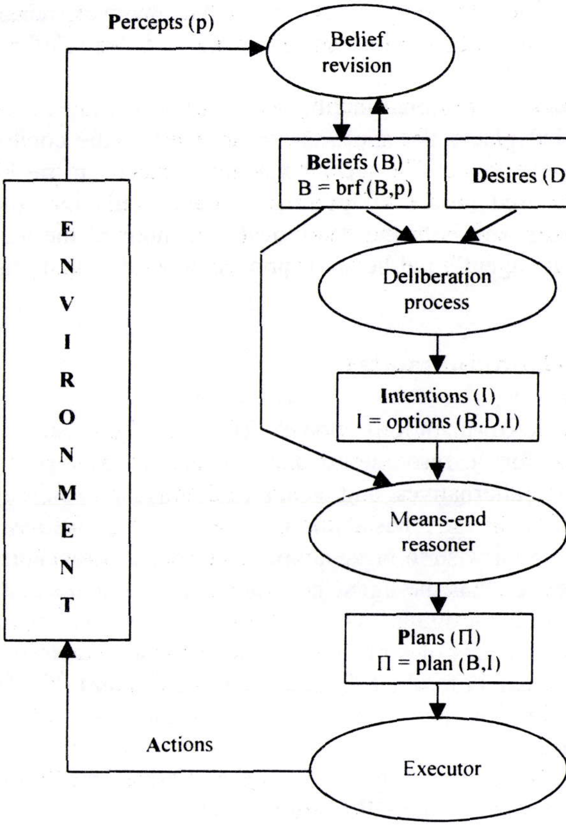


**Figure 1**: A BDI agent model

**Agent control loop**

B = $B_0$; D = $D_0$; I = $I_0$; { $B_0$, $I_0$ may be empty }
**while** true **do**
    get next percept p;
    B = brf (B, p);
    I = options (D, B, I);
    Π = plan (B, I);
    **while not** (empty (Π) **or** succeeded (I,B) or impossible (I, B)) **do**
        α= head (Π); Π = tail (Π);
        execute (α);
        get next percept p;
        B = brf (B, p);
        **if not** sound (Π, I, B) **then** Π = plan (B, I);
    □
□

**Figure 2:** Agent behavior according to the BDI model

As the agent is autonomous, it will continuously execute an agent control loop, as shown by the algorithm in Figure 2. The algorithm is adapted from the one in [13] with minor modifications and changes. Depending on its desires and beliefs, the agent will develop a plan of actions to achieve its desires and will execute the actions in the environment. The agent's desires may be either consistent or contradictory. In case of contradictory desires, the agent has to select a consistent subset of desires, which becomes the set of goals to be achieved (in case of consistent desires, these may be equated to goals).

We have developed an *emotional modeling architecture*, called **BDE (Belief-Desires-Emotions),** which is built upon the BDI architecture and includes the agent's emotions and emotional state. The BDE model emphasizes emotional reasoning in the context of the other forms of reasoning specific to BDI. A BDE agent has sets of **beliefs** ($B$), **desires** ($D$), and **intentions** ($I$), but has also a set of **emotions** ($E_i$) and a set of **emotional states** ($E$). An event may generate one or more emotions that are to be integrated into an emotional state. The set of emotional states may be the same as the set of individual emotions if there is only one emotion generated by each event. Moreover. $E_i$ is also equal to $E$ if the integration process will select one prevailing emotion among several competing ones.

A BDE agent combines its BDI reasoning with emotional reasoning based on the events perceived from the environment. As in the BDI model, a BDE agent has a set of percepts by means of which it recognizes an event, which can be either a state change or an action occurrence. The percept obtained from the outside world may cause a change of beliefs, leading to a belief revision process, but may also generate emotions according to the emotion-generation process, as Figure 3 shows.

An agent recognizes the event and matches it against its desires, beliefs, and intentions to obtain a *qualitative appraisal of the event*. The mechanism is somehow similar to the one of the OCC theory but it is the BDI components of the agents that are used to appraise the event instead of the agent's goals, standards, and attitudes, as in OCC. Let us consider the following example. Agent Adina has the desire to attend CASYS'03 and the belief that she will actually present a paper to the conference. A new event occurs in the environment: Adina has broken her leg. The event is appraised against the agent's desires and beliefs and the resulting emotion is sadness, as the agent realizes her desire is no longer achievable.

As stated above, emotions are integrated into an emotional state. The emotional state may then influence the beliefs, desires, and intentions of an agent but also its plans. An emotional state may lead to belief revision but also to **desire revision**, a reasoning process not present in the BDI model. For example, if agent John is not keeping a promise to Edgar, Edgar will be disappointed, he will change the belief about John being a trustful friend and he will also change his desire to tell John a big secret. One should note that, without the emotion of disappointment, there is no reason for the agent Edgar to change his desire.

Emotions are decayed over time trying thus to mimic human emotional behavior. In the same time, some strong emotions may have a long lasting effect: even if the emotion has fainted it may be remembered if the emotion was very strong. To account for this other feature of human emotional behavior, we have introduced the **Emotional Memory** ($E_M$) component, which memorizes particular powerful emotions experienced over time. Emotion generation will thus be based not only on the current perceived event but also on previously experienced emotions.

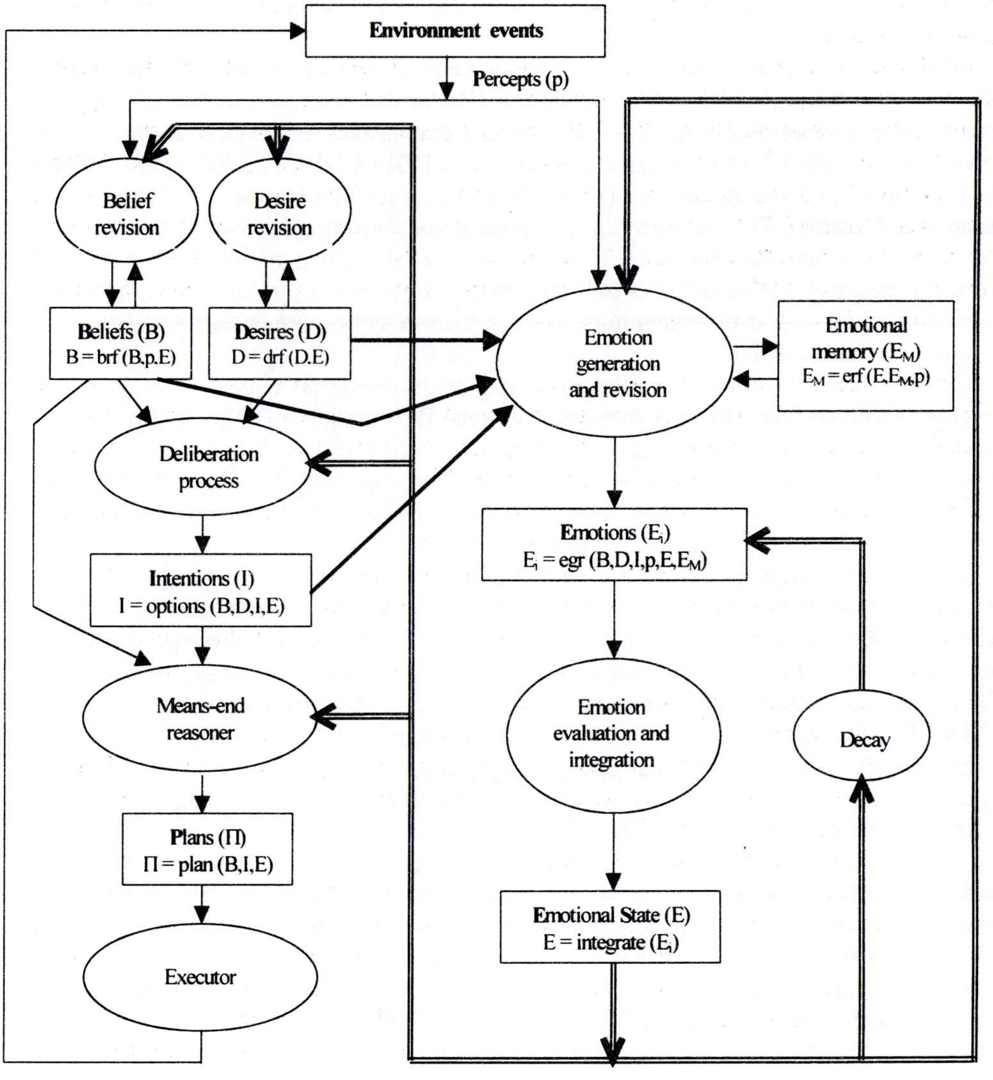


**Figure 3**: A BDE agent model.

The emotional reasoning components of a **BDE agent** (Figure 3) are denoted by the following functions:

- *egr*: $B \times D \times I \times p \times E \times E_M \rightarrow E_i$ is the emotion generation function, with $p$ the set of percepts;
- *integrate*: $E_i \rightarrow E$ is the function to integrate different emotions (if several emotions are generated by the same event) into one emotional state; as explained previously, $E_i$ may be identical to $E$ depending on the integration model;
- *erf*: $E \times E_M \times p \rightarrow E_M$ is the function which retrieves from the Emotional Memory, if the case, a previously experienced emotion for the same percept $p$.

The *brf*, *options*, and *plan* functions from the basic BDI model are similar except for the fact that they have one extra argument, namely the emotional state obtained as a result of the agent's emotional reasoning.

There is one supplementary function to account for the desire revision process depending on the emotional state: *drf*: $D \times E \rightarrow D$. The algorithm in Figure 4 describes the control loop of the BDE agent.

**BDE Agent control loop**

$B = B_0$; $D = D_0$; $I = I_0$; $E = E_0$; { $B_0$, $I_0$ and $E_0$ may be empty }
**while** true **do**
    get next percept p;
    $E_i$ = egr (B, D, I, p, E, EM);
    E = integrate ($E_i$);
    B = brf (B, p, E);
    **if not** empty (E) **then** D = drf (D, e);
    I = options (D, B, I, E);
    Π = plan (B, I, E);
    **while not** (empty (Π) **or** succeeded (I,B) **or** impossible (I, B)) **do**
        α= head (Π); Π = tail (Π);
        execute (α);
        get next percept p;
        $E_i$ = egr (B, D, I, p, E, EM);
        E = integrate ($E_i$);
        B = brf (B, p, E);
        **if** E affects D **then** D = drf (D, E);
        **if** E affects I **or** changed(D) **then** I = options (D, B, I, E);
        **if not** sound (Π, B) **or** E affects Π **or** changed(I) **then** Π = plan (B, I, E);

**Figure 4**: Agent behavior according to the BDE model

## 4 Emotion generation, integration and influence

The emotional reasoning functions described in the previous sections are generic functions. Depending on a particular event, one or several components of the BDE architecture will be considered when appraising the event to generate the emotion. They will be called *emotion eliciting conditions* (*EEC*). The *influence of the emotional state on behavior* (*IEB*), although indicated in all of the reasoning functions (as a function argument), will be present or not depending on the particular emotion. We shall exemplify this process by analyzing a set of basic emotions that we have selected: satisfaction, joy, hope, sadness, anger, fear, and disappointment.

**Satisfaction**

*EEC*: the event confirms a belief (or several beliefs) or the event is an action, which achieves an intention (or several intentions), but the desire that determined the intention is not yet accomplished.

*IEB*: satisfaction will generate belief revision if a belief is confirmed or, in case an intention is achieved, may lead to new options or to new plans (replanning).

**Joy**

*EEC*: the event achieves one or more desires (or goals – see discussion about desires and goals in the previous section).

*IEB*: joy will lead to desire revision, and, if necessary, new options and new plans.

**Hope**

*EEC*: the event adds a belief which removes contradictions between desires or a belief which facilitates finding a plan to satisfy a desire or goal.

*IEB*: hope will lead to new options (a new deliberation process) and to new plans.

**Sadness**

*EEC*: the event makes unachievable one of the agent's desires or makes an intention no longer possible to be achieved due to a state change (the agent that caused the event, namely the state change, is unknown)

*IEB*: sadness may cause desire revision, new options, and new plans.

**Anger**

*EEC*: the event is the action of another agent and it makes unachievable one of the agent's most valued desires or intentions.

*IEB*: anger will cause new options and replanning, but may also lead to belief revision and desire revision.

**Fear**

*EEC*: the event represents a threat to one of agent's desire or to an intention as regarding intention completion; it does not actually make the desire or intention impossible, but endangers it.

*IEB*: fear will cause inspecting new options (selection of a new set of intentions) and may also cause investigation of new alternate plans.

**Disappointment**

*EEC*: the event contradicts one of the agent's beliefs or the event removes a belief according to which an intention was chosen.

*IEB*: disappointment may cause belief revision, new options and new plans.

Satisfaction and hope determine the agent to anticipate the successful achievement of the corresponding desire, while fear makes it anticipate a possible failure.

Emotion integration (see Figure 4) comprises two separate processes: evaluate the intensity of the emotion and integrate several emotions into a single emotional state if an event's appraisal leads to different emotions simultaneously. Emotions may have different intensities. The OCC model [3] proposes eight intensity variables for each emotion type, for example goal importance, unexpectedness, arousal, and others. We believe that a computational model of emotions should consider fewer variables to evaluate the emotion intensity. We propose *three variables to evaluate the intensity of an emotion*: desire's preference, unexpectedness of the event, and desire's plausibility.

In the BDI model, desires have, usually, associated preferences. The agent uses these preferences to select a consistent subset of desires from its contradictory desires, identifying thus its goals. Therefore, in our BDE model, the goal importance variable of the OCC model may be represented by the **desire's preference** of the agent. Unexpectedness is related to how unexpected en event is and has a strong connection with the anticipatory capabilities of the agent. The BDE agent has a set of beliefs, which may be contradicted by some events. Some of these beliefs may have been used as the basis for intention selection, following the *options* function, and for plan synthesis, following the *plan* function. If one of these beliefs is contradicted by an event we consider the unexpectedness of this event higher that the one of an event contradicting a belief which was not used in the current control loop. Moreover, the unexpectedness of an event is correlated with the number of contradicted beliefs on which intentions and plans were based. A larger number of contradicted beliefs will assign a higher level of **unexpectedness** to the event. A third measure that may be used to evaluate the intensity of an emotion is the probability of goal attainment, as in the Emile system [7], where the agent's goals have associated probabilities. In our model, we propose a third intensity variable, which is somehow related to goals' probability, namely the **desire's plausibility**. Desire or goal's plausibility depends on the number of alternate options and plans an agent has for achieving that goal. An event affecting intentions or plans to achieve a goal with a higher plausibility will generate a less intense emotion, as the agent believes there are alternate ways to achieve that goal.

The result of the event appraisal process may lead to one or more emotions. Psychologists have explored the link and influences between different emotions that may occur in the same time. For example, the Cathaxis [12] model is based on an energy metaphor where different appraisals act as energy elicitors that excite or inhibit different emotional states, and decay over time. The model suggests that some emotions may be excited by some other emotional states (e.g., joy excites hope) and inhibited by others (e.g., joy may modestly inhibit fear). In our model we propose to use a set of simple rules with conditions that test the generated emotions and their intensity, and conclude an overall emotional state. Such a rule-based approach has two advantages: it is straightforward and it may be tailored to the particular set of emotions and selected psychological model. The particular set of rules is currently under study with the contribution of psychologists interested in the field.

## 5 Conclusions

In both human and artificial agents' activity, anticipation of behavior plays a crucial role. To this end, behavior is based on models of themselves, of the other agents in the environment, and of the environment. For humans, the predictive model is not always a precise and clearly defined one, and has many attributes that are difficult to model or quantify. Emotion is such a complex concept and certainly influences behavior and the predictive capability of humans. In an attempt to simulate part of such a behavior, we have developed the Belief-Desire-Emotion emotional model of reasoning. Starting from the OCC theory of emotion generation in human beings, we have adapted the theory to the BDI model by anchoring the event appraisal process in the beliefs, desires and intentions of an agent. We have described the different stages of the emotion generation process and captured the way emotions may influence both the theoretical and practical reasoning of an agent.

We have also selected a set of basic emotions and we have shown how these particular emotions fit in the proposed model by analyzing both their activation based on external events and their specific influence on the agent's behavior. Based on our approach, the proposed set of emotions may be easily extended to include the analysis of a richer set of emotions.

Future research issues that should be addressed are: evaluate the proposed model in different contexts, extend the model to include the representation of other BDE agents and the capacity of an agent to reason about other agents' emotions, include a learning algorithm to learn desire's plausibility, and tailor communication between agents according to their emotional state. We also consider integrating a more detailed analysis of the emotional influence on the planning process, based on the work presented in [7]. Additionally, the model should be enhanced with synthetic personalities of the agents, allowing thus emotional modeling to become more accurate and specific.

Despite of the on-going debate about the possibility of building anticipatory behaviour in artificial systems, we consider emotional intelligent agents as a next step towards creating artificial systems with strong predictive features, based on both rationality and emotions. Such agents are reasonable by being also "irrational" as the agents use both rational and emotional reasoning. Modeling emotional reasoning interplayed with other types of reasoning gives a better approximation of the way humans behave and solve problems while interacting with other humans.

## References

[1] Ekdahl B. (2000) Agents as anticipatory systems. In the 4th World Multiconference on Systemic, Cybernetics and Informatics, SCI 2000, Orlando, Florida.

[2] Rao A.S. and Georgeff M.P. (1991) Modeling Rational Agents Within a BDI-architecture. Proc. of Knowledge Representation and Reasoning (KR&R-91). Edited by R. Fikes and E. Sandwall, published by Morgan Kaufman, pp. 473-484.

[3] Ortony A., Clore G., and Collins A.(1988) The Cognitive Structure of Emotions. Cambridge University Press, Cambridge, UK.

[4] Reilly W.S. (1996) Believable social and emotional agents. PhD Thesis, Carnegie Mellon University, Pittsburgh, PA.

[5] Elliott C. and Brzezinki J. (1998) Autonomous agents as synthetic characters. AI Magazine, AAAI Press, pp.13-30.

[6] El-Nasr M. S., Ioerger T. R., and Yen J. (2000) A Web of emotions. Texas A&M University Technical Report.

[7] Gratch J. (2000) Modeling the interplay between emotion and decision-making. The 9th Conference on Computer Generated Forces and Behavioural representation, Orlando, Florida. Online at http://www.ict.usc.edu/publications/cgf00.pdf

[8] Eliasmith C. (1999) Dictionary of the philosophy of mind. Online at http://www.artsci.wustl.edu/~philos/MindDict/

[9] Bratman M. E. (1990) What is intention? In Intentions in Communication. Edited by P.R. Cohen, J.L. Morgan, and M.E. Pollak. Published by MIT Press, Cambridge, MA, pp.15-32.

[10] Dubois D. M. Introduction to Computing Anticipatory Systems. International Journal of Computing Anticipatory Systems. Edited by Daniel M.Dubois. Published by CHAOS, Liège, Belgium, 1998, vol. 2, pp. 3-14.

[11] Dubois D.M. (2000) Review of Incursive, Hyperincursive and Anticipatory Systems - Foundation of Anticipation in Electromagnetism. Computing Anticipatory Systems: CASYS'99 - Third International Conference. Edited by Daniel M. Dubois, Published by The American Institute of Physics, AIP Conference Proceedings 517, pp. 3-30.

[12] Dubois D.M. (2003) Mathematical Foundations of Discrete and Functional Systems with Strong and Weak Anticipations. Anticipatory Behavior in Adaptive Learning Systems - Foundations, Theories, and Systems Edited by Butz M.V., Sigaud O. and Gérard P. Published by Springer. LNAI 2684, pp.110-132.

[13] Wooldridge M. (2000) An Introduction to MultiAgent Systems. John Wiley & Sons, England.

[14] J. Velasquez. (1998) When Robots Weep: Emotional Memories and Decision-Making. Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI-98), Madison, WI, MIT/AAAI Press, pp.70-75.