Affective Agents for Empathic Interactions

Flávio Soares Corrêa da Silva and Ary Fagundes Bressane Neto

Department of Computer Science, Institute of Mathematics and Statistics,
University of São Paulo, Rua do Matão, 1010, São Paulo, Brazil
{fcs,bressane}@ime.usp.br

Abstract. In the present work we develop an experimental setting to evaluate the influence of affect – more precisely, of simulated coherent fluctuations in mood and emotional states – in the construction of empathy in synthetic characters for interactive digital entertainment. Our goal is to evaluate whether the impression of interacting with human-like agents can be more relevant than purely utilitarian metrics for user preferences to interact with systems in certain situations. We have built affective agents to play computer games against human users, and assessed empirically the extent to which users consider more engaging to play against the affective agents than to play against agents that are easy to beat.

Keywords: Affective Computing, Digital Entertainment, Artificial Intelligence, Intelligent Agents.

1 Introduction

Research in Artificial Intelligence can be broadly organised as normative or descriptive, with very little work lying in the intersection of these two categories. Normative research has a focus on the construction of models to optimise the performance of artificial systems, thus determining how (artificial) intelligent agents ought to reason, instead of how (natural) agents actually reason [16].

Descriptive models of artificial intelligence, in turn, focus on the reconstruction of observed human performance in engineered intelligent artifacts, in order to build scientific theories of human rationality [6]. Descriptive theories and models can be useful in at least two situations, namely (i) when one is interested in understanding human behaviour, which, although not optimised, can be predicted and, therefore, either guided or exploited in constructive ways [1], and (ii) when one is interested in the construction of empathic agents, with which human users can interact pleasantly and naturally [11].

In the present work we focus on descriptive models of intelligent agents. More specifically, we focus on the construction of empathic agents for interaction with human users, in the context of digital entertainment. We consider computer games in which the user plays against the computer, and build synthetic opponents whose performance mimics that of human opponents, this way building pleasant and satisfactory experiences to users.

We have explored the utilisation of an agent architecture enriched with a model for personality and mood fluctuations, to build synthetic opponents whose behaviour matches that of human players [10]. Preliminary experiments have indicated that this approach to build computer games does indeed create novel possibilities for the game industry, to build appealing new products for consumers. Moreover, the same technology can be used to build empathic agents for a variety of applications, such as smart homes for the routine care of elderly and disabled people, conversational agents for computer-based education and for information presentation.

In section 2 we briefly discuss related work that has been inspirational to our work. In section 3 we introduce our enriched agent architecture. In section 4 we introduce the two experiments that we have built in order to empirically assess the practical utilisation of the enriched agent architecture. In section 5 we discuss some preliminary empirical results. Finally, in section 6 we present some conclusions and proposed future work.

2   Related Work

Several projects have been developed aiming at the construction of believable agents for interaction with human users. Some of these projects have been particularly influential on the development of the present work.

The ALMA Project [5] is connected to the Virtual Human Project, whose aim is the development of a system for distance learning, in which users interact with synthetic characters in order to obtain instructive material. The synthetic characters are built using a model of personality, mood and emotional states. The main focus on ALMA is on physiological manifestations, although it also comprises a few features related to cognitive states, such as the selection of strategies for dialogues.

Making Them Remember [7] aims at the development of agents with personality and emotions, whose memory of past events changes according to emotional states. Its results have been tested in a synthetic character applied to distance learning.

The Governor of Poker[1] is a commercial product, in which users can play Texas Hold'em Poker against the computer. The computer reacts based on emotional states generated by its performance in the games. For example, if the user makes a bluff and the computer falls on it, then the computer can lose emotional control and play badly for the next few rounds.

Additional references to related initiatives can be found in the literature, as reviewed in previous publications from the authors [10]. We have furthered those initiatives, focusing on the cognitive consequences of changes in affective states.

3   Agents with Personality and Mood

We have extended a standard BDI architecture for intelligent agents, to include models of personality, mood and emotions, in such way that these models influence cognitive properties of the agent, namely the way it generates plans and its capability to perceive and interpret external events, as well as to store and retrieve past events

---

[1] http://www.youdagames.com/Governor-of-Poker-game-info-2818

from the memory. Emotional states also determine physiological responses from the agent, such as (in our experiments) facial expressions.

As a base architecture, we have adopted Jason[2], which is an open source framework to build BDI intelligent agents. We have extended Jason with an Affective Module and three filters, namely a Perception Filter, a Memory Filter and a Planning Filter. The Affective Model stores (static) information about an agent's personality, as well as (dynamic) information about the agent's mood and emotional state, which is affected by external events as well as by its success to build adequate courses of action in response to events. The filters are influenced by the Affective Model and determine the extent to which, respectively, the agent is going to be able to perceive and interpret external events, store and retrieve information from its memory and build effective plans.

A detailed presentation of the extended agent architecture can be found in previous publications from the authors [10]. In the next paragraphs we present in further detail the Affective Module and each of the three filters (namely, the Perception, Memory and Planning filters) used in the proposed architecture for affective agents.

3.1 The Affective Module

The Affective Module influences the three filters, and consequently how the agent perceives external events, stores and retrieves information from the memory and generates plans. Internally, the module contains and interconnects models for personality, mood and emotion. The interconnections are implemented using the model of Mehrabian [9], which expresses the relationships between personality, mood and emotion using the parameters of the Big Five Personality Factors [8], the Pleasure-Arousal-Dominance (PAD) model for mood [9] and the OCC model of emotion [14]. This model has been used with success in previous research initiatives [5, 7].

3.2 Personality

In our model, personality is assumed to be stable for an agent. We have adopted the Big Five Personality Factors [8] to model personality, which characterises the personality of an individual according to five variables (Openness, Conscientiousness, Extroversion, Agreeableness and Neuroticism). Each variable can assume a real value in the interval [-1; 1], such that -1 means the extreme opposite of a variable and 1 means the maximum value for a variable. For example, -1 for Extroversion identifies an extremely introverted individual, and 1 for the same variable identifies an extremely extroverted individual. Hence, personality can be portrayed using a vector $P = [O, C, E, A, N], O, C, E, A, N \in [-1,1]$.

3.3 Mood

Mood is a state that results from the cumulative effects of emotions over an initial state defined by personality [4]. In our work, we have adopted the model of Mehrabian [9], which characterises mood based on three variables (namely, Pleasure,

---

Arousal and Dominance). Similarly to the model for personality, each variable can assume a value in the interval [-1, 1]. Stereotypical mood states are characterised by extreme values for each variable, and hence we have eight stereotypical values. Considering that mood is represented using a vector $M = [P, A, D], P, A, D \in [-1,1]$, the stereotypical values are [1, 1, 1] = Exuberant, [1, 1, -1] = Dependent, [1, -1, 1] = Relaxed, [1, -1, -1] = Docile, [-1, -1, -1] = Bored, [-1, -1, 1] = Disdainful, [-1, 1, -1] = Anxious, and [-1, 1, 1] = Hostile.

The initial mood state for an agent has been empirically related to personality [9]:

$$\begin{cases} P = 0{,}21P[E] + 0{,}59P[A] + 0{,}19P[N] \\ A = 0{,}15P[O] + 0{,}30P[A] - 0.57P[N] \\ D = 0{,}25P[O] + 0{,}17P[C] + 0{,}60P[E] - 0{,}32P[A] \end{cases}$$

3.4 Emotion

Emotion is a state of immediate effect and short duration. It is activated by external events, actions or specific objects, and influences physiological manifestations like facial expressions. We have adopted the OCC model of emotions [14], which describes 22 types of emotions (Admiration, Anger, Disappointment, Distress, Fear, Fear-confirmed, Gloating, Gratification, Gratitude, Happy-for, Hate, Hope, Joy, Love, Pity, Pride, Relief, Remorse, Reproach, Resentment, Satisfaction, Shame), and a set of rules that determine how these emotions vary according to external events and objects, as well as agent actions. In our model, we have implemented the OCC model based on a vector $E = [e_1, \dots, e_{22}], e_i \in [0,1]$.

3.5 The Relations between Personality, Mood and Emotion

Personality is defined for an agent as a combination of values for the variables O, C, E, A, N. Given the personality of the agent, mood and emotion are inter-related as follows.

Initially, mood is obtained from the set of equations above. After that, the agent iteratively obtains information about the environment and external events, and updates mood and emotions, using the equations empirically defined in [7]. In these equations, E_s represents the emotional state, E_a represents the appraised emotional state, M_{cur} represents the current mood state and the α_{ij} represent constant values that characterise the cross-influences between emotion and mood [5, 9]:

Emotion update

$$E_s = E_s + filter(E_a, M_{cur})$$

$$filter(E_a, M_{cur}) = E_a + \frac{\sum_{i=1}^{22}\sum_{j=1}^{3} \alpha_{ij} * m_j}{\sum_{i=1}^{22}\sum_{j=1}^{3} \alpha_{ij}}$$

Mood update

$$M_{cur} = M_{cur} + UpdateMood(E_s)$$

$$UpdateMood(E_s) = \sum_{i=1}^{22}\sum_{j=1}^{3} e_i * \alpha_{ij}$$

3.6 Perception

Empirical research has indicated that the intensity of perception of stimuli in individuals varies according to their mood [13, 12]. We have extended Jason, so that perception now receives two new attributes: a vector indicating which emotions are associated to this perception and a vector that determines the intensity of each emotion. We have also implemented the Perception Filter, in order to select perceptions following two criteria: the perception associated to an intensive emotion and whether the result of the appraised emotions (i.e. whether the emotions are predominantly positive or negative) is consistent with the current affective state of the agent.

For the first case we have calculated a summation of the intensity of emotions and, if this value is greater than the value of the minimum intensity (I_{min}) defined by the user, the perception is selected, otherwise it is discarded. For the second case, assessment is made on the number of positive and negative emotions caused by perception. The dominant (positive or negative) type is compared to the current affective state of the agent (in all three filters, the current affective state of the agent is represented by the value of the variable P(leasure) in the mood vector). If the values are congruent, then the perception is selected, otherwise it is discarded.

3.7 Memory

According to Bower [2], there are four theories that match affective states and cognitive processes: (1) memories are more easily retrieved when an individual has the same mood as that of when they were stored; (2) the current affective state influences reasoning and interpretation of events; (3) memories with affective state similar to the current one are easily retrieved; and (4) memories associated to more intense emotions are more promptly retrieved. In our architecture, the memory of an agent is stored as its belief base. Affective states and their intensities can provide a context to influence the retrieval of information from memory, hence we have built a Memory Filter that selects information related to the current emotional state of the agent from the belief base. Stored information is tagged with the emotions associated to it, their intensity, and the current affective state of the agent.

Information is retrieved from the belief base according to the following criteria: the intensity of the emotion associated to the information must be greater than the minimum intensity defined by the user (I_{min}); the type (positive or negative) of emotion associated with the stored information must be congruent with the current affective state; and the affective state at the moment when the information was stored must be congruent with the current affective state.

3.8 Planning

The influence of affective states in plan formation is based, in our architecture, on the model of Somatic Markers [3]: before an individual makes a rational decision, the brain reduces the number of options based on somatic aspects of recent past experiences. In that model, each plan is tagged with a positive or negative marker, depending on the current emotional state of the agent. The Plan Filter makes a pre-selection of the potentially applicable plans, rejecting those that are connected to

negative markers, as well as updating the tags of all applicable plans for future consideration. Somatic markers implement an adaptive behavior, which selects appropriate plans in such way as to increase the efficiency of plan selection based on congruence with the emotional state of the agent.

4   Experiment Design

We have built two experiments, in order to assess the extent to which synthetic characters endowed with personality, mood and emotion can build more engaging experiences to users than characters which do not feature these attributes.

Both experiments are online computer games, in which a user plays against the machine. The first experiment is a version of the Memory Game, and the second experiment is a version of the Texas Hold'em Poker Game.

Our goal was to assess the influence of fluctuations in mood and emotional states in the pleasure generated by playing the game in users, and whether this influence could overcome the pleasure that resulted from winning in the game.

4.1   Experiment 1: The Memory Game

The Memory Game is a two-player game, in which a deck of cards containing printed images has pairs of identical cards. The cards are shuffled and then positioned on a table facing down, so that the players cannot see their images. Each player flips up two cards.

If the cards are identical, then the player removes the cards from the game and scores one point, otherwise the player flips the cards down again and passes the turn to the other player. The goal of the game is to score the maximum number of points. Since both players can see the images on the cards that are flipped up, then both players can memorise the location of certain cards as the game goes on.

The construction of a computer game for the Memory Game, in which a human player can play against the computer, can be challenging, despite the simple rules of the game, because it can be difficult to build an engaging synthetic opponent for this game. For example, a perfectly rational opponent with perfect memory of past events can be too hard to beat, and an opponent that flips cards randomly can be too easy to beat. Even though the user feels pleasure as a result from beating the computer, our conjecture was that relevant pleasure could also result from the sense of control that results from causing variations on the cognitive and affective state of the opponent, hence users could prefer to play against the affective character, even though other, nonaffective characters could be easier to beat.

In our implementation, we have used 16 pairs of cards, as shown in Figure 1. Each image portrays a character showing one different emotion in a different intensity. Emotions are classified as negative or positive, depending on the emotional states that they evoke, and the intensity of the emotions induced by each image grows from left to right in Figure 1. A screenshot of the game is shown in Figure 2.
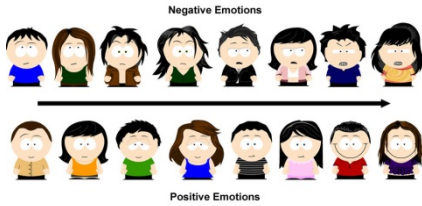
Fig. 1. Characters in the Memory Game
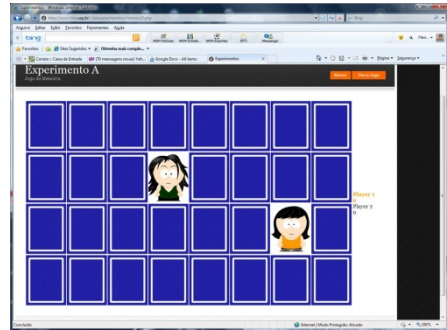
Fig. 2. The Memory Game – Screenshot

There are three possible opponents against which the user can play: (1) Random Player, which picks pairs of cards at random, (2) Perfectly Rational Player, which has infinite memory and never makes a mistake, and (3) Affective Player, which employs the architecture described in the previous sections to change behaviour depending on mood and emotional states – more specifically, the mood of the Affective Player changes depending on how successful it has been in recent previous rounds. Successful rounds increase the sense of control of the agent, which memorises the cards portraying positive emotions better than those portraying negative emotions, and vice-versa when unsuccessful rounds occur.

Simulations have shown that this strategy for the Affective Player – enhanced by random mistakes whose likelihood varies, also depending on the mood of the agent – mimics very effectively the behaviour of a human player. Our focus in the present work is whether the pleasure that results from playing against a human-like player can overcome the pleasure of winning the game.

4.2 Experiment 2: Kuhn Poker

The Kuhn Poker is a simplified version of a two-player poker, in which the deck of cards has only three cards: Jack (J), Queen (Q) and King (K), (K is the best card and J is the worst one). Each opponent starts the game with 10 chips that can be used to bet. To simplify the presentation of the Kuhn Poker, we name the players A and B. Initially, each player bets one chip and receives one card. A can then choose between raise – i.e. betting an additional chip – or check – i.e. asking to see the cards. Upon raise, B can choose between call – i.e. betting one matching chip and showing cards to see who wins the round – or fold – i.e. giving up. Upon check, B can choose between check – i.e. showing cards to see who wins the round – or raise – i.e. betting one extra chip and leaving A with the options call or fold. The winner always takes all chips on the table. In case of a draw, the chips remain on the table for the next round.

As in more complex versions of poker, in this version the players can bluff. In this experiment, we update mood and emotional state of a synthetic opponent depending on its performance on recent rounds. The skills to play, as well as to deceive the human opponent using its facial expressions vary depending on the mood.

Following the rules of the Kuhn Poker, we have limited each game to three rounds, and built the affect-influenced strategy for the synthetic opponent using tables as the sample tables presented in Tables 1, 2, 3 and 4. In these tables, the first column shows the card in the hand of the synthetic opponent, the second column shows its mood and the third column shows the corresponding action. The mood of the synthetic opponent varies according to previous rounds, in such way that positive mood and emotional states result from winning chips, and negative states result from loosing chips. The tables shown in this article feature the initial strategy of the player, which is updated according to mood and emotional states. The facial expression of the player mimics that of an intermediate player: at some points during the game it reflects the emotional state of the player, and at other points it portrays what the player wants to convey – for example, a bluffing expression of joy.

Table 1. 1st Round Strategy

Card	Mood	Action
J	Exuberant, Relaxed, Disdainful, Hostile	Check
J	Dependent, Docile	Check
J	Bored, Anxious	Raise
Q	Exuberant, Relaxed, Disdainful, Hostile	Check
Q	Dependent, Docile, Bored, Anxious	Raise
K	Exuberant, Relaxed, Disdainful, Hostile	Raise
K	Dependent, Docile	Raise
K	Bored, Anxious	Check

Table 2. 2nd Round Strategy - after a Check

Card	Mood	Action
J	Exuberant, Relaxed, Disdainful, Hostile	Check
J	Dependent, Docile	Check
J	Bored, Anxious	Raise
Q	Exuberant, Relaxed, Disdainful, Hostile	Raise
Q	Dependent, Docile, Bored, Anxious	Check
K	Exuberant, Relaxed, Disdainful, Hostile	Raise
K	Dependent, Docile	Raise
K	Bored, Anxious	Check

Table 3. 2nd Round Strategy - after a Raise.

Card	Mood	Action
J	All	Fold
Q	Exuberant, Relaxed, Disdainful, Hostile	Call
Q	Dependent, Docile, Bored, Anxious	Fold
K	All	Call

Table 4. 3rd Round Strategy

Card	Mood	Action
J	All	Fold
Q	Exuberant, Relaxed, Disdainful, Hostile	Call
Q	Dependent, Docile, Bored, Anxious	Fold
K	All	Call

In order to complete our experimental setting, we have created two other players. Our set of three alternative synthetic opponents is as follows: (1) the Random Player selects at random its bets. It also shows in its facial expression the quality of its hand (it looks happy when it has a K and it looks sad when it has a J). This opponent is very easy to beat; (2) the Affective Player behaves as described in the previous paragraphs. Its facial expressions show how it feels with respect to the game as a whole, but do not reveal how good its hand is; and (3) the Skilled Player, which

employs an optimised strategy to play the game, as presented in [15]. Its facial expression remains unchanged (the "poker face") during the whole game. This opponent is very hard to beat.

Our implementation of the Kuhn Poker is depicted in Figure 3.
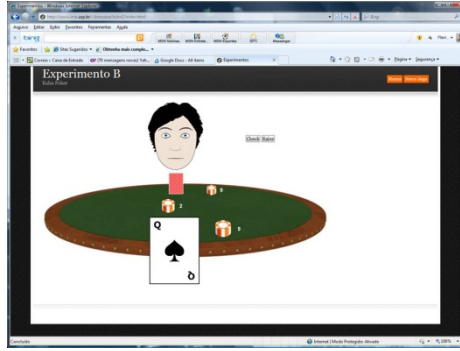

Fig. 3. The Kuhn Poker - Screenshot

5   Experimental Results

Our experimental setting, in both experiments, was as follows: we have asked volunteers to play against the three types of opponents, without explaining the difference between them. After playing a few matches against all three opponents, volunteers were asked to answer the following questions: (1) against which opponent did you enjoy playing most? (2) which opponent did you consider the hardest one to beat? (3) which opponent would you consider the most human-like?

Classical economic theories suggest that users should always prefer to play against opponents which are easy to beat. As we show in the next paragraphs and diagrams, however, this is not observed empirically. We have asked 23 participants to play the Memory Game, and 16 participants to play the Kuhn Poker.

In the Memory Game, 74% of the participants considered that the affective player was the most human-like, and 69% of the participants enjoyed playing against this player most. In the Kuhn Poker, 56% of the participants considered that the affective player was the most human-like, and 56% of the participants enjoyed playing against this player most. We have observed the behaviour of the participants while playing, and additional observations, although pointwise, are worth mentioning. In particular, one of the participants, while playing against the affective character in the Kuhn Poker experiment, started swearing against the computer when bad results weer obtained or when she suspected that the computer could be bluffing, and cheering when good results were obtained. Similar behaviour was observed with a different participant while playing against the affective character in the Memory Game.

The answers to each of the three questions are plotted in Figure 4 for the Memory Game, and in Figure 5 for the Kuhn Poker.
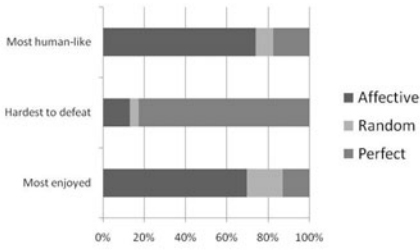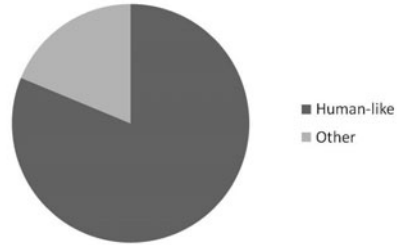
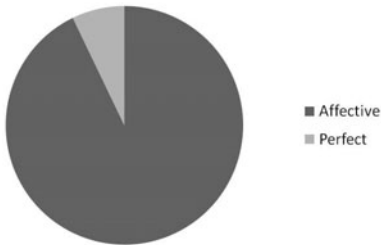**Fig. 4.** The Memory Game - Survey Results
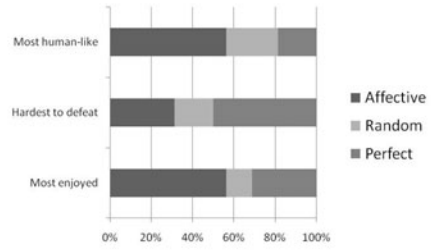
Fig. 5. The Kuhn Poker - Survey Results

Fig. 6. The Memory Game – Preference for Human-like Opponents

Fig. 7. The Kuhn Poker – Preference for Humanlike Opponents

Fig. 8. The Memory Game – Specific Preference for Affective Human-like Opponents

Fig. 9. The Kuhn Poker – Specific Preference for Affective Human-like Opponents

Some participants have considered that players which were not our affective player were the most human-like. Even in these cases, the majority of participants stated that they preferred to play against the opponent which, according to their evaluation, was the most human-like. 61% of the participants stated that they enjoyed most to play against the most human-like opponent in the Memory Game, and 81% of the participants stated the same thing in the Kuhn Poker. Considering only these participants, 93% have stated their preference for the affective opponent in the Memory Game and 62% have stated the same in the Kuhn Poker.

These results are plotted in Figure 6 for the Memory Game and Figure 7 for the Kuhn Poker regarding the preference for human-like opponents, and in Figure 8 for the Memory Game and Figure 9 for the Kuhn Poker regarding the preference for the affective opponent in the restricted universe of participants who had already stated their preference for human-like opponents.

6 Conclusion and Future Work

Preliminary empirical results have confirmed that human-like behaviour can be relevant to build engaging interactions with users, depending of course on the nature of the system under development. Interactive systems that are approached for clearcut reasons and desired goals (e.g. ATMs) are more likely to be expected to behave as computers instead of as simulated humans, whereas systems which are more conversational in nature (e.g. computer games, tutoring systems, museum guides, etc.) can be expected to assist the interactions, and behave in general as simulated humans. This feature can improve the empathy of information systems, and in order to build it we need well founded simulation models, based on which the cognitive responses of the system vary according to personality, mood and emotional states.

This feature, when applied in relevant systems, seems to be particularly relevant for beginning users: in our experiments, we have also asked the participants to declare how often they played computer games, and out of the participants who did not declare that they preferred the affective opponent over the other two alternatives, all participants were heavy game players for the Kuhn Poker, and all but one were heavy players for the Memory Game.

The perception of a human-like opponent induced by fluctuations in mood and emotional states seems to be more effective in simpler systems. In the case of the Kuhn Poker, which is a more sophisticated system, a competitive behaviour from the opponent seemed to suffice to convey the perception of human-like behaviour in a significant number of cases, whereas in the Memory Game the use of mood and emotions to influence the actions of the opponent seemed to be more relevant to convey the same perception.

Future work shall be devoted to further analysis of the situations in which affective agents can be most relevant for intelligent interactive systems. We shall build experiments with larger groups of participants, so that we can analyse the influence of factors such as gender, age and academic degrees to discriminate further our results. We shall also build experiments with different sorts of systems, e.g. systems containing embodied conversational agents to provide users with requested information.

References

1. Ariely, D.: Predictably Irrational, Revised and Expanded Edition: The Hidden Forces That Shape Our Decisions. Harper Perennial, USA (2010)

2. Bower, G.H.: Mood and memory. American Psychology, 129–148 (1981)
3. Damasio, A.R.: Descartes' Error: Emotion, Reason, and the Human Brain. Harper Perennial, USA (1994)
4. Ekman, P., Davidson, R.J.: On emotion, mood, and related affective constructs. In: The Nature of Emotion, UK. Oxford University Press, Oxford (1994)
5. Gebhard, P.: Alma: a layered model of affect. In: Proceedings of the 4th International Conference of Autonomous Agents and Multiagent Systems (2005)
6. Ito, J.Y., Pynadath, D.V., Marsella, S.C.: Self-deceptive decision making: normative and descriptive insights. In: Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2009), Hungary (May 2009)
7. Kasap, Z., Moussa, M.B., Chaudri, P., Magenant-Thalmann, N.: Making them remember: emotional virtual characters with memory. In: Proceedings of the IEEE Computer Graphics and Applications (2009)
8. McCrae, R.R.: An introduction to the five-factor model and its applications. Journal of Personality, 175–215 (1992)
9. Mehrabian, A.: Pleasure-arousal-dominance: a general framework for describing and measuring individual differences in temperament. Current Psychology, 261–292 (1995)
10. Bressane Neto, A.F., Corrêa da Silva, F.S.: On the Construction of Synthetic Characters with Personality and Emotion. In: da Rocha Costa, A.C., Vicari, R.M., Tonidandel, F. (eds.) SBIA 2010. LNCS(LNAI), vol. 6404, pp. 102–111. Springer, Heidelberg (2010)
11. Nguyen, H., Masthoff, J.: Designing empathic computers: the effect of multimodal empathic feedback using animated agent. In: Proceedings of the 4th International Conference on Persuasive Technology, USA. ACM Press, New York (2009)
12. Niedenthal, P.M., Halberstadt, J.B., Margolin, J., Innes-Ker, A.H.: Emotional state and the detection of change in facial expression of emotion. European Journal of Social Psychology, 211–222 (2000)
13. Niedenthal, P.M., Setterlund, M.B.: Emotion congruence in perception. Personality and Social Psychology Bulletin, 401–411 (1994)
14. Ortony, A., Clore, G.L., Collins, A.: The cognitive structure of emotion. Cambridge University Press, UK (1988)
15. Southey, F., Hoehn, B., Holte, R.C.: Effective short-term opponent exploitation in simplified poker. Machine Learning 74, 159–189 (2009)
16. Thagard, P.: Why cognitive science needs philosophy and vice versa. Topics in Cognitive Science 1, 237–254 (2009)