

# Social Capabilities for Autonomous Virtual Characters

Christopher Peters<sup>1</sup>, Catherine Pelachaud<sup>1</sup>, Elisabetta Bevacqua<sup>1</sup>, Magalie Ochs<sup>1,2</sup>,  
Nicolas Ech Chafai<sup>1,2</sup>, and Maurizio Mancini<sup>1</sup>

<sup>1</sup> IUT de Montreuil, Université de Paris 8  
<http://www.iut.univ-paris8.fr/greta>

<sup>2</sup> France Télécom, Division R&D, France

**Abstract.** In this paper we describe our work toward the creation of affective multimodal virtual characters endowed with communicative and other socially significant capabilities. While much work in modern game AI has focused on issues such as path-finding and squad-level AI, more highly-detailed behaviour for small groups of interacting game characters has been lacking in the literature. We apply our extensive background knowledge from the domain of Embodied Conversational Agents (ECA's) to the task of creating autonomous humanoid game characters capable of displaying believable and expressive social behaviours. We believe that in order to reach our aim, we ought to endow the agents with human-like capabilities. The agents should be able to sense the virtual world and recognise the intentions of the other agents, to engage in and maintain a conversation, to have the ability to be engaged, to communicate verbally and non-verbally and to display expressive behaviours. In this paper we describe our methodology and framework.

## 1 Introduction

Interactive computer games have become an increasingly useful, if not challenging domain, for the testing and application of human-level artificial intelligence techniques [21]. In this paper, we address capabilities for enhancing the autonomous behaviour of humanoid virtual agents for populating game worlds. We foresee these entities, endowed with human-like capabilities, playing an ever more important role in the next generation of computer games as they become increasingly sophisticated. Although virtual humans have a human-like appearance, this is not their main asset: what really matters in game-play terms is their ability to resonate with the player through naturalistic communicative and emotional capabilities. These agents are able to sense the environment they are placed in, they can interact with other agents, they can talk to them, react to their utterances and other events in the world. Virtual characters populate video games, chat rooms, and have been particularly prominent in massively multiplayer online role playing games (MMORPG's). A common difficulty in each of these spheres is the animation of the characters: Pre-scripted animation is simply not adequate for truly interactive gaming. Autonomy in terms of deciding which action to take and which behaviour to show is a must for plausibly animated characters. In games that contain a social aspect, such as virtual communities in MMORPG's, an essential role of virtual characters is to communicate with other characters, both user and computer controlled. Communication involves complex processes. It is fundamentally multimodal.

Nonverbal behaviours do have a major role in the interaction between people. They have several functions: they may be used to process the environment and others; they may indicate our mental and emotional state; they may be used as signals to be read by others; they may also be used to organise our thoughts. In our work, we are interested in the first 3 functions. The endowment of a virtual character with these functions requires for it to be able to: (1) generate multimodal communicative behaviour; (2) have emotions and display expressive behaviours; (3) perceive the environment and the other's gaze and behaviours; (4) manage the conversation by showing feedback.

Because virtual environments are often inhabited by multiple virtual characters, variability in agents should occur not only in their geometric appearance, but also in how they react to events and behave: their emotion models and expressive and multimodal behaviour models should accept different parameters so that they act differently.

## 2 Emotion

Emotion is a crucial element in any gaming experience: virtual characters who are able to display emotions are more likely to be able to invoke an emotional reaction in the user and thus add to the users gaming experience. Despite this, sophisticated emotion models have not yet been incorporated into virtual characters. One way to add an emotional dimension to a game is to create an affective virtual agent, one that utilises a computational model of emotion. The primary purpose of such a model is to control expressive subsystems, such as a facial expression module, in order to convey a plausible impression of an emotional episode in the virtual character to the user. In this Section, we first describe emotional capabilities of an affective agent; we then propose a method to realise such an agent.

First of all, an affective virtual agent should be able to adopt an emotional behaviour to convey its emotions. Three emotional processes have to be simulated [34]: the elicitation of emotion, expression of emotion and experience of emotion. These allow an agent to identify the circumstances in which an emotion is triggered, display emotions and manifest emotional influence on their reasoning and actions, respectively. An affective agent should also be able to represent its environment in terms of its own emotions and those of other characters [8]. The integration of these emotional capabilities enable one to improve several characteristics of virtual agent's:

- **Believability**: the expression of emotion enables a virtual agent to create an illusion of life and increase the user's engagement [4].
- **Emotional relationship**: to create an emotional relationship between characters (and users), the emotional capabilities presented here are necessary [44].
- **Autonomy**: emotions can serve the purpose of selecting appropriate actions or memorising and retrieving information [6].
- **Emotion-orientation**: the affective virtual agent chooses actions according to the character's or user's emotions that it wants elicit, creating emotion-oriented games.

*Appraisal theory* [43] is one way in which we may give an agent the ability to identify the emotional meaning of a situation. According to this theory, emotions are elicited by a subjective interpretation of an event, which depends both on situational and cultural

factors and on a particular individual's features. Major determinants in emotion elicitation include the person's beliefs and goals. The interpretation of an event corresponds to the appraisal of a set of variables called appraisal variables. Particular combinations of these variables lead to emotion elicitation (using, in many cases, the OCC model of emotion [27]). Appraisal variables are often hard-coded [11, 39] or based on the agent's plan [15]. In contrast, we aim to create a domain-independent model of emotion elicitation that enables the virtual agent to assess its own emotions and those of other characters without having prior access to their plans.

We propose to represent emotion elicited-events based on a BDI approach [38]. The mental state of a BDI agent is composed of mental attitudes such as beliefs and intentions. It corresponds to its cognitive representation of the world at a given instant. It includes a representation of the events perceived in the environment. Accordingly, an occurred emotion elicited-event is also represented through mental attitudes. Emotion elicited-events can then be represented by a particular mental state. Based on the OCC model [27], we have described values of appraisal variables through beliefs and intentions. For instance, a desirable event corresponds to the belief that an occurred event enables an agent to achieve one of its goals. According to appraisal theory, we have represented emotion elicited-events by combinations of values of appraisal variables; that is, by a combination of mental attitudes.

From this formalisation of emotion elicited-events, a virtual agent can identify in real-time its own emotions triggered by a situation. By knowing other characters intentions, their emotions can also be identified. More details on this computational model of emotion elicitation and associated model of facial expressions of emotions can be found in [26]. Finally, existing games such as *Façade* [24] and *The Sims* [2] are marked out for using computational models of emotion elicitation.

### 3 Communicative behaviours

In social settings, the study of proper communicative behaviours are of importance for believable and engrossing game characters. Humans communicate in a rich variety of ways through multiple channels: through our choice of words, facial expressions, body postures, gaze, gestures... Non-verbal behaviours accompany the flow of speech and are synchronised at the verbal level, punctuating accented phonemic segments and pauses. They may have several communicative functions [35, 7]. They are used to control the flow of conversation; that is they help in regulating the exchange of speaking turns, keeping the floor or asking for it. Actions such as smiling, raising the eyebrows, and wrinkling the nose often co-occur with a verbal message. They may substitute for a word or string of words, or emphasise what is being said. Gestures may indicate a point of interest in space or describe an object. Facial expressions are the primary channel to express emotion. They can also express attitude toward one own speech (such as irony) or toward the interlocutor (like showing submission). Non-verbal behaviour do not occur randomly, but rather are synchronised to one's own speech, or to the speech of others (see for example, [20, 42]). Raised eyebrows go along emphasised words [10]; the stroke of a gesture, that is the most forceful part of a gesture, happens also on emphasised word or just before [25]. Most of the gesturing happens when speaking.

Hands and faces come to rest when speech ends. Because of the inherent diversity of behaviour, implementing a computation model is challenging for virtual characters and requires consideration of the communicative function of different behaviours.

### 3.1 Communicative functions

In our model, we follow the taxonomy of communicative functions proposed in [35]. A communicative function is defined as a pair (meaning, signal). Each function may be associated with different signals; that is, for a given meaning, there may be several ways to communicate it. For example, the meaning ‘emphasis’ (of a word) may co-occur with a raise eyebrow, a head nod, a combination of both signals, or even a beat gesture. Vice versa, a same signal may be used to convey different meanings; e.g. a raise eyebrow may be a sign of surprise, of emphasis, or even of suggestion [35]:

1. information about speaker’s beliefs: behaviours that provide information on the speaker’s beliefs such as the degree of certainty regarding what he is talking about.
2. information about speaker’s intentions: the speaker may provide information on his goal through, for example, his choice of performative or the focus of his sentence.
3. information about speaker’s affective state: the speaker may show his emotion state through particular facial expressions.
4. metacognitive information about speaker’s mental state: the speaker may try to remember or recall an information.

In our model, the agent’s behaviour is synchronised with its speech [28]. It is consistent with the meaning of the sentences it pronounces. To control the agent’s behaviour we are using a representation language, called ‘Affective Presentation Markup Language’ (APML) where the tags of this language are these communicative functions [9]. In APML, text to be spoken by the agent is annotated with tags denoting emotion and interaction information, for example:

```
<turn-allocation type=“take turn”> <performative type=“greet”>  
Good Morning, Angela.  
</performative> </turn-allocation> <affective type=“happy”>
```

Spoken text is synthesised automatically for playback with generated face and body animations.

### 3.2 Behaviour Variation

A further area of importance when animating the communicative behaviour of humanoid characters is to ensure that they exhibit individualised behaviour. That is, their behaviour should be consistent with their personality, mood, emotional state and with their individual characteristics.

The realisation of such behaviours is particularly needed in videogames where human players interact with other human or artificial players using virtual representations (like in *MMORPGs*). Without these, virtual worlds would seem to be populated by androids



**Fig. 1.** Left: greeting gesture that gets instantiated from the APMML example in Section 3. APMML is converted in facial and body animation parameters that are in turn used to animate the agent. Right and inset: the agent in the real-time Torque game environment. Among other capabilities, it is also capable of real-time text-to-speech synthesis and facial animation.

and zombies. Synthetic characters have to differ not only in how they look (a popular research area in contemporary videogames) but also in the way they behave [18]. As discussed in Sections 2 and 3, virtual agents' visible behaviour is determined by their emotive state and goals of conversation. In real-life situations there are 2 further considerations: first, any given behaviour, for example a gesture / facial expression, will be performed in a slightly different way by two individuals even if the internal state causing the gesture / facial expression are (theoretically) exactly the same; second, given the same emotive state and goals of conversation, two individuals will choose different behaviours (gestures / facial expressions) to perform. That is, it is not enough to fix an emotive state and a goal of conversation to be sure of the final behaviour of a person. There are other factors, related to individuality, that will determine the behaviour. We will consider in more detail two aspects of our model for individualised agents: the expressivity model and the multimodal model.

**Expressivity model** Many researchers (see, for example, Wallbott and Scherer [45], Pollick [37]) have investigated human motion characteristics and encoded them into categories such as slow / fast, small / expansive, weak / energetic, small / large, unpleasant / pleasant. We define the *expressivity* of behaviour as the "quality" of the communicated information through the execution of some physical behaviour. Starting from the results reported in [45], we have defined and implemented expressivity [17] as a set of parameters that affect the gesture (performed by arms or head) quality of execution speed of arms / head, spatial volume taken by the arms / head, energy and fluidity of arms / head movement, number of repetitions of the same gesture. Thus, the same gestures or facial expressions are performed in a physically different way depending on

the emotional state of the agent, something with great promise for generating variable character behaviours.

**Multimodal behaviour generation model** A social agent should be able to interact with the user through all the modalities involved in human-human communication: speech, gestures, gaze, facial expressions, body movements and body posture [29, 41]. We implemented a mechanism to choose between these different available modalities during conversation. Our model is based on a hierarchy of modalities in which each modality is associated with a value representing its degree of preference. For example, an agent with the tendency to use gestures during communication will have a higher value associated with the gesture modality. Hierarchy values can be determined from cultural studies (for example, Italian people communicate with an higher number of gestures than other people), contextual factors (for example, during a conversation in a noisy place, people may tend to use more gestures than usual) or idiosyncratic behaviours (for example, a person may have an inborn tendency to always use head movements during speech).

During the communication process our agent will choose different signals (gestures, facial expressions, etc) depending on the modalities hierarchy, the meaning of what the agent says and the goals of conversation: emotional states are mainly conveyed through facial expressions while descriptions of objects are usually achieved through hand gesturing.

## 4 Perception and attention

The ability to sense information from the environment in a manner congruent with real humans is an important prerequisite for plausible agent behaviour: as the old programming adage goes, “garbage in, garbage out”. An agent that senses the virtual environment in a manner different to a human is doomed to behave differently to a human, no matter how sophisticated its cognitive capabilities. Furthermore, games such as *Thief* and *Half-Life* have demonstrated how even simplified (when compared with real-life) visual and acoustic models can enhance game-play by providing agents with a better repertoire of perceptual capabilities [22].

We provide agents with real-time synthetic vision, attention and memory capabilities (see [31]). All three capabilities interact with each other in order to collect information and orient the agents senses with respect to the environment. Our visual sensing module is monocular with multiple resolutions; it operates in a snapshot manner by taking frequent updates of the visible region of the scene from the point of view of the agent. Renderings are taken at each perceptual update: a full-scene rendering and a false-colored rendering. During false-colored rendering, objects are preassigned unique false-color values and rendered in a simplified manner using these values as colors and later matched back up with their database counterparts. Although the synthetic vision approach is more computationally intensive than geometric methods alone [13], it forms an important basis when accurate visual attention capabilities are desired in both an object- and spatially-based view-dependent manner; this approach inherently

accounts for lighting conditions and other rendering effects, details which may have significant behavioural consequences.

The purpose of the visual attention module is to choose a set of visible objects or locations in the scene for enhanced processing by the agent and is based on a popular real-time model from computational neuroscience [19]. We use it with virtual scenes by passing full scene renderings from the synthetic vision module in order to compute the saliency map: a 2D grey-scale spatial ranking of the scene in terms of areas that ‘pop-out’ and are thought to attract visual attention. A memory system stores uncertainty levels for each object to ensure the focus of attention moves throughout the scene; values are combined into an uncertainty map, with which the saliency map is modulated in order to provide the final attention map. Artificial regions of interest are enumerated from the attention map and are used to create artificial scan-paths for the agent.

In a gaming environment, these capabilities provide autonomous and human-like looking behaviours and help drive more subtle movements, such as eye saccades and blinking. These systems serve as a basis for implementing higher-level cognitive and behavioural mechanisms: for example, we have used these capabilities, along with theory of mind, to control autonomous interaction initiation between agents (see Section 5.1).

#### **4.1 Theory of Mind**

In parallel to agent perception, attention and memory, we have implemented a special pathway for social perception processing [32] by incorporating a perceptual theory of mind model based on evolutionary psychology and supported by neurophysiological evidence [30]. In a virtual community it is vital that avatars and agents act in a social manner and pay special attention, as well as be seen paying such attention, to the actions of other goal-directed entities. In our model, an intentionality detection (ID) mechanism filters perceived agents into this special pathway and segments them into eye, head and body subparts. A direction of attention detector (DAD) computes the direction of the eyes, head and body of the other agent with respect to the self. Orientations are merged into an *attention level* metric through a subpart weighting process, which is boosted if mutual attention exists between the agents, as signalled by a mutual attention mechanism (MAM). Results are stored in short-term memory. An agent may integrate these metrics over a time interval on demand in order to formulate theories of whether the other has seen it, whether the other has seen this agent see it and the interest that the other agent appears to have in this agent. This system is useful for generating emergent gaze behaviours between agents based on the perception of the actions of the other.

## **5 Communication Management**

Communication is by nature bi-directional. The barrier between speaker and listener can be quite fuzzy. Both send signals and adapt continuously to each other. We often talk about sender and receiver rather than speaker and listener [36]. Communication involves a start and an end. Both events are tightly linked to conventional behaviours where gaze, in particular, and other behaviours such as head movement and gesture

have a major role. Through gaze we may indicate we wish to start a conversation; by stopping gesturing we are finishing our speaking turn or even an interaction; on the other hand we may signal we wish to take the speaking turn by starting to gesticulate. A holistic communication management must have the ability to handle communication initiation and maintenance of conversation by assessment of feedback obtained from the other.

### 5.1 Start of interaction

An important but surprisingly rarely investigated topic when considering interaction between humans and/or virtual agents is how interactions begin in the first place. Interaction initiation behaviours such as gaze orienting, waving and name utterances are of great significance for adding realism to virtual communities where visual social interaction behaviours are desired between agents and avatars [18].

In real environments, people must routinely deploy their senses in order to scan for possible contacts, must gain the attention of one they wish to interact with, must signal interest in communicating both verbally and non-verbally and must seek the cooperation, as well as gauge the willingness, of the other to reciprocate in conversation. For a human viewer, in behavioural plausibility terms, sometimes it is not the destination that counts (i.e. conversation), so much as getting there (i.e. the acts involved in initiating the conversation). We implement an automatic non-verbal interaction initiation system based primarily on direction of attention and theory of mind (Section 4) to drive a hierarchical finite state machine between various interactional states, such as monitoring the environment, grabbing the attention of the other and gauging the reaction of the other. Thus, unlike other systems, agent behaviour is based not only on the goals of the agent, but also on its theory of the intentions of the other in order to avoid the social embarrassment of engaging in conversation with an unwilling participant [14]. Our evaluation of conversation initiation behaviours seems to suggest that users may be as sensitive to subtle non-verbal social cues in the virtual environment as in the real situation [33].

### 5.2 Feedback

During a conversation, two flows of information are established between speaker and listener, one concerning the effective content of the speaker's speech and the other regarding the listener's reaction to this speech. Interlocutors are expected to inform the speaker about the success of communication, providing signals showing whether they are engaged and interested in the conversation, whether they understand and agree and furthermore which emotional and attitudinal reactions are elicited in them by the speaker's speech [1, 36]. These kind of signals are called *feedback* and without them a communication becomes difficult and frustrating since the speaker is not able to understand if the listener is following the conversation or also how he is being affected by the interaction. Therefore, if we want to implement virtual agents capable of human-like interaction skills, we have to take into account the listener's behaviour: it is a fundamental aspect of human-human communication.



Video games where users control agents embedded in conversational settings should not lack a model that generates appropriate feedback behaviours. To be able to display believable behaviour, virtual characters should not freeze after they stop speaking to give the turn to another agent, but they should keep on moving, emitting feedback signals to show that the interaction is still going on. *Façade* [24] is an example of this new generation of games where users control virtual characters that can communicate with each other. This game illustrates one type of listener feedback: for instance, if the player uses words or sentences that the game does not understand, the virtual agents may look confused or appalled, or humour the player with awkward small talk.

Feedback strongly depends on the context; usually the listener provides signals according to what the speaker says in order to demonstrate specific information about his reaction to the speech content or to provoke a particular effect on the speaker, for example: the listener decides to stare at the speaker to show disbelief or surprise expecting a confirmation by the speaker. We call this kind of feedback *cognitive feedback*. However, the listener also emits signals without thinking, unconsciously [1]. In fact, during a conversation, the listener makes a lot of behavioural decisions in such a short time that he is not even aware of them. He reacts instinctively to the speaker's behaviour or speech, generating feedback signals unconsciously. This kind of feedback is called *reactive feedback*. Consequently a single feedback model is not enough; two computational models are needed, respectively a cognitive model and a reactive model. The latter model, however, is quite hard to implement and to make real-time. To elaborate reasoned reactions from a listener, one must have access to not only the extrapolated speech content, but also information about the listener's personality. For this reason, we are implementing just the reactive model.

Since the instinctive listener's feedback is often elicited by the speaker's behaviour, a set of rules can be defined [23, 12]. For example, from a corpus of data, Maatman derived a list of rules useful to predict when feedback can occur according to the speaker's actions. Backchannel continuers (like head nods, verbal responses) appear at a pitch variation in the speaker's voice; frowns, body movements and gaze shifts are produced when the speaker shows uncertainty; facial expressions, postural and gaze shifts are provided to reflect those made by the speaker (mimicry).

### 5.3 Interaction Maintenance

As we mentioned, communication management involves an agent behaving in a specific way in order to engage in a conversation and to take into account the behaviour of the listener that reacts to the speaker's discourse. Another aspect of an interaction is to consider the gestures that the speaker produces. In video games, based on conversational interaction with a user, the use of gestures may enhance the communication with a specific capability.

Gestures are the result of the speaker's ideation [5] and help speaker to formulate his thoughts. Gestures create a link with the speakers thoughts and have a pragmatic value in the sense that they illustrate his communicative efforts. To maintain a listener into a conversational interaction, we need to study this pragmatic dimension of gestures. Some previous work conducted by [16, 3] collected objective data from an eye tracker to study which gestures elicit focal attention: when a speaker uses a gesture to depict

or to point at an object, the listener usually tends to gaze at this gesture or toward the object. Thus, gestures are part of the communication process.

In our works, we collected subjective data from a corpus of conversational interactions in traditional animations. This corpus provides data regarding the expressive performance of animated characters. Considering the expressive parameters of gestures (fluidity, power, spatial expansion, and repetitivity), we observe that these parameters are not only the expression of the character's personality or emotional state, but also reflect the pragmatic intention of the speaker. This latter property is, at least partially, fulfilled through the modulations of gesture expressivity over time, which suggests some type of relationship (relation of similarity or relation of contrast) between the elements of the verbal or non-verbal utterance. If we want to create games in which agents are able to maintain the user's interest in a conversational interaction, we should take into account the pragmatic value of gestures and consider the rhetorical relations that they are able to underline in discourse.

## 6 Conclusion

This paper has presented a number of autonomous agent capabilities that we are confident will contribute to the next generation of real-time game characters. It should be mentioned that our work focuses on low-level high-detail interaction scenarios between a small number of agents, in contrast to simulations of larger groups or crowds, (see for example [40]). One challenge is to extend our models beyond scenarios involving only two interactants, to those involving more agents. This is challenging, not only because of the computational complexity involved, but also because of the predominance of existing theory for low number of interactants, making formulation of appropriate models difficult. Furthermore, we also focus here solely on autonomous agents: an important challenge for application to computer games is how to mix such autonomous control methods with director controlled methods.

We are in the process of integrating these capabilities into a system for real-time agents in a computer game environment. Already, using the *Torque* game engine (<http://www.garagegames.com>), we have integrated perception, attention and interaction capabilities into an agent as outlined in Sections 4 and 5.1 (see Fig. 1). We are in the process of transferring and integrating further capabilities, including all of those mentioned here, as well as investigating perceptual processing distributed between multiple computers or implemented on the GPU.

## 7 Acknowledgements

This work has been partially funded by the Network of Excellence Humaine (Human-Machine Interaction Network on Emotion), IST-2002-2.3.1.6 / Contract no. 507422 (<http://emotion-research.net/>).

## References

1. J. Allwood, J. Nivre, and E. Ahlson. On the semantics and pragmatics of linguistic feedback. *Semantics*, 9(1), 1993.

2. Electronic Arts. The sims.
3. G. Barrier, J. Caelen, and B. Meillon. La visibilité des gestes: Paramètres directionnels, intentionnalité du signe et attribution de pertinence. In *Workshop Français sur les Agents Conversationnels Animés*, pages 113–123, Grenoble, France, 2005.
4. J. Bates. The role of emotion in believable agents. *Commun. ACM*, 37(7):122–125, 1994.
5. G. Calbris. *L'expression gestuelle de la pensée d'un homme politique*. CNRS Editions, 2003.
6. L. Cañamero. *Emotions in Humans and Artifacts*, chapter Designing Emotions for Activity Selection in Autonomous Agents, pages 115–148. MIT Press, 2003.
7. N. Chovil. Social determinants of facial displays. *Journal of Nonverbal Behavior*, 15(3):141–154, Fall 1991.
8. C. Crawford. *On Game Design*. New Riders Games, 2003.
9. B. DeCarolis, C. Pelachaud, I. Poggi, and M. Steedman. Apml, a mark-up language for believable behavior generation. In Helmut Prendinger and Mitsuru Ishizuka, editors, *Life-Like Characters*, Cognitive Technologies. Springer, 2004.
10. P. Ekman. About brows: Emotional and conversational signals. In M. von Cranach, K. Foppa, W. Lepenies, and D. Ploog, editors, *Human ethology: Claims and limits of a new discipline: contributions to the Colloquium*, pages 169–248. Cambridge University Press, Cambridge, England; New-York, 1979.
11. C. Elliot. *The Affective Reasoner: A process model of emotions in a multi-agent system*. Computer science, Northwestern University, 1992.
12. D. Friedman and M. Gillies. Teaching virtual characters to use body language. In *Intelligent Virtual Agents*, Lecture Notes in Artificial Intelligence. Springer-Verlag, 2005.
13. J. Funge. *Artificial Intelligence For Computer Games: An Introduction*. AK Peters, Wellesley, MA, July 2004.
14. E. Goffman. *Behaviour in public places: notes on the social order of gatherings*. The Free Press, New York, 1963.
15. J. Gratch. Emile: Marshalling passions in training and education. In *The Fourth International Conference on Autonomous Agents*, pages 325–332, Barcelona, Catalonia, Spain, 2000.
16. M. Gullberg and K. Holmqvist. Keeping an eye on gestures: Visual perception of gestures in face-to-face communication. *Pragmatics and Cognition*, 7:35–63, 1999.
17. Bjoern Hartmann, Maurizio Mancini, and Catherine Pelachaud. Towards affective agent action: Modelling expressive ECA gestures. In *Proceedings of the IUI Workshop on Affective Interaction*, San Diego, CA, January 2005.
18. K. Isbister. *Better Game Characters by Design: A Psychological Approach*. Elsevier Science and Technology Books, 2006.
19. L. Itti. *Models of Bottom-Up and Top-Down Visual Attention*. PhD thesis, California Institute of Technology, Jan 2000.
20. A. Kendon. Movement coordination in social interaction: Some examples described. In S. Weitz, editor, *Nonverbal Communication*. Oxford University Press, 1974.
21. J. E. Laird and M. van Lent. Human-level AI's killer application: Interactive computer games. In *AAAI/IAAI*, pages 1171–1178, 2000.
22. T. Leonard. Building an ai sensory system - examining the design of thief - the dark project. In *Proceedings of Game Developers Conference 2003*, San Francisco, California, 2003. CMP Game Media Group.
23. R. M. Maatman, J. Gratch, and S. Marsella. Natural behavior of a listening agent. In *5th International Conference on Interactive Virtual Agents*. Kos, Greece, 2005.
24. M. Mateas and A. Stern. Façade: An experiment in building a fully-realized interactive drama. In *Game Developers Conference, Game Design track*, March 2003.
25. D. McNeill. *Hand and Mind: What Gestures Reveal about Thought*. University of Chicago, 1992.

26. Magalie Ochs, Radoslaw Niewiadomski, Catherine Pelachaud, and David Sadek. Intelligent expressions of emotions. In Jianhua Tao, Tieniu Tan, and Rosalind W. Picard, editors, *The 1st International Conference on Affective Computing and Intelligent Interaction (ACII'05)*, pages p. 707–714, Beijing, China, 2005. Springer.
27. A. Ortony, G.L. Clore, and A. Collins. *The cognitive structure of emotions*. Cambridge University Press, 1988.
28. C. Pelachaud and M. Bilvi. Computational model of believable conversational agents. In Marc-Philippe Huget, editor, *Communication in Multiagent Systems*, volume 2650 of *Lecture Notes in Computer Science*, pages 300–317. Springer-Verlag, 2003.
29. C. Pelachaud, V. Carofiglio, B. De Carolis, and F. de Rosis. Embodied contextual agent in information delivering application. In *First International Joint Conference on Autonomous Agents & Multi-Agent Systems (AAMAS)*, Bologna, Italy, July 2002.
30. D.I. Perrett and N.J. Emery. Understanding the intentions of others from visual signals: neurophysiological evidence. *Current Psychology of Cognition*, 13:683–694, 1994.
31. C. Peters. *Bottom-Up Visual Attention for Autonomous Virtual Human Animation*. PhD thesis, Department of Computer Science, Trinity College Dublin, 2004.
32. C. Peters. Direction of attention perception for conversation initiation in virtual environments. In *International Working Conference on Intelligent Virtual Agents*, pages 215–228, Kos, Greece, September 2005.
33. C. Peters. Evaluating perception of interaction initiation in virtual environments using humanoid agents. In *Proceedings of the 17th European Conference on Artificial Intelligence*, pages 46–50, Riva Del Garda, Italy, August 2006.
34. R. Picard. *Affective Computing*. MIT Press, 1997.
35. I. Poggi. Mind markers. In N. Trigo M. Rector, I. Poggi, editor, *Gestures. Meaning and use*. University Fernando Pessoa Press, Oporto, Portugal, 2003.
36. I. Poggi. Backchannel: from humans to embodied agents. In *Conversational Informatics for Supporting Social Intelligence and Interaction - Situational and Environmental Information Enforcing Involvement in Conversation workshop in AISB'05*. University of Hertfordshire, Hatfield, England, 2005.
37. F. E. Pollick. The features people use to recognize human movement style. In A. Camurri and G. Volpe, editors, *Gesture-Based Communication in Human-Computer Interaction - GW 2003*, number 2915 in LNAI, pages 10–19. Springer, 2004.
38. A. S. Rao and M.P. Georgeff. Modeling rational agents within a bdi-architecture. In *Proceedings of International Conference on Principles of Knowledge Representation and Reasoning (KR)*, pages 473–484, San Mateo, CA, USA, 1991.
39. S. Reilly. *Believable Social and Emotional Agents*. Computer science, University of Carnegie Mellon, 1996.
40. C. W. Reynolds. Interaction with groups of autonomous characters. In *Proceedings of Game Developers Conference 2000*, pages 449–460, San Francisco, California, 2000. CMP Game Media Group.
41. Z. Ruttkay and C. Pelachaud. Exercises of style for virtual humans. In *Symposium of the AISB'02 Convention*, volume Volume Animating Expressive Characters for Social Interactions, London, 2002.
42. A.E. Schefflen. The significance of posture in communication systems. *Psychiatry*, 27, 1964.
43. K. Scherer. *Introduction to Social Psychology: A European perspective*, chapter Emotion. Oxford, 2000.
44. A. Stern. *Emotions in Humans and Artifacts*, chapter Creating emotional relationships with virtual characters. MIT Press, 2003.
45. H. G. Wallbott and K. R. Scherer. Cues and channels in emotion recognition. *Journal of Personality and Social Psychology*, 51(4):690–699, 1986.