Building Expression into Virtual Characters

Article · January 2006

Source: OAI

7 authors, including:


Vinoba Vinayagamoorthy
BBC
26 PUBLICATIONS **675** CITATIONS


Marco Gillies
Goldsmiths, University of London
69 PUBLICATIONS **532** CITATIONS

Anthony Steed
University College London
131 PUBLICATIONS **2,873** CITATIONS


Xueni Pan
University College London
26 PUBLICATIONS **184** CITATIONS

Building Expression into Virtual Characters

V. Vinayagamoorthy[1], M. Gillies[1], A. Steed[1], E. Tanguy[2], X. Pan[1], C. Loscos[1], and M. Slater[1,3]

[1]Department of Computer Science, University College London, London, UK
[2]Department of Computer Science, University of Bath, Bath, UK
[3]ICREA-Universitat Politècnica de Catalunya, Barcelona, Spain

Abstract

Virtual characters are an important part of many 3D graphical simulations. In entertainment or training applications, virtual characters might be one of the main mechanisms for creating and developing content and scenarios. In such applications the user may need to interact with a number of different characters that need to invoke specific responses in the user, so that the user interprets the scenario in the way that the designer intended. Whilst representations of virtual characters have come a long way in recent years, interactive virtual characters tend to be a bit "wooden" with respect to their perceived behaviour. In this STAR we give an overview of work on expressive virtual characters. In particular, we assume that a virtual character representation is already available, and we describe a variety of models and methods that are used to give the characters more "depth" so that they are less wooden and more plausible. We cover models of individual characters' emotion and personality, models of interpersonal behaviour and methods for generating expression.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Virtual Reality

1. Introduction

Virtual characters are an essential part of many interactive 3D graphics simulations. They are used within human-computer interaction in order to invoke peoples' automatic responses to the human form and behaviour, and thereby achieve a kind of empathic interaction that would otherwise be difficult. Virtual characters might convey part of the story as a computer game, or they might act as commanders in a military simulations. The designers of such systems are not just adding virtual humans to make the scene realistic, but the characters are the primary mechanism to create content, back-story and mood within the simulated scenarios. It is important that the virtual characters are plausible within the context of the scenario, and of course this means that the virtual humans must be perceived by the user to be an authentic part of that scenario.

The aim is to be able to construct virtual humans to which people respond as if they were real, even though at a high level of mental processing, the user knows fully well that they are not. However, this is difficult to achieve in practice. Whilst in the last few years a lot of work has been done on the visual plausibility of characters, once the user is allowed to interact with the virtual characters, the illusion of believability can disappear rapidly. Often the behaviour of the characters can look "wooden", exhibiting perhaps, a lack of variety in response, a lack of emotional range or a lack of adaptation to the users' attitude towards the character. In real-life humans are all unique and have limitless variety in their response, whereas virtual characters often have little "depth".

In this STAR we give an outline of models and methods that are used to create expressive virtual characters through the creation and representation of a psychological state. This psychological state might evolve based on the personality of the characters and their interactions with each other and/or the user. Such an expressive character should be more believable because it should reflect the simulated situation more like the user would expect them to. To create this type of behaviour in virtual characters, a concrete understanding of human behaviour is needed. This understanding can be provided by theoretical models from psychology and its related disciplines.

This STAR is targeted at two audiences. The first is developers and researchers who have a character simulation system and want to improve a scenarios' believability. For this audience we have outlined implemented systems that model an aspect of expressiveness. The second audience is those wanting to push the boundaries of expressive characters, either within a system or through research. For this audience we have included relevant theoretical background and we have commented on which models might have advantages over more commonly implemented models. Throughout the STAR we assume that a virtual character with full body pose and facial expression control systems is already available hence we do not discuss how to create geometry and surface details for characters. We focus on real-time interactive humanoid characters in this STAR. Much of the material is also relevant for off-line characters, though in that case, much more emphasis is put on skill of the animator.

The STAR is structured as follows. Firstly in the following section we give more motivation for the use of virtual characters. This will help us see what requirements are in demand for virtual characters and some of the ways in which characters are (or can be) evaluated. Then in section 3, we describe nonverbal communication as the means by which a character achieves expression. In effect, nonverbal communication is the medium through which the virtual character will convey a psychological state. Though we stress the importance of achieving synchronicity between verbal and nonverbal communication, we do not deal with verbal communication here since this would require a STAR of its own.

Sections 4 and 5 focus on models of emotion and personality. Emotion is modelled in order to give an immediate context to a characters' behaviour whilst personality is modelled to give that emotion itself a context in terms of the disposition and interactions of the character. Personality thus gives mechanisms for emotion to change over time. Section 6 deals with interpersonal factors. Although a character can have a personality and express emotions, these are of limited use unless these reflect the behaviour of other characters (or the user). A virtual character that must interact with another must be "socially intelligent", for instance, it must undertake proper turn-taking behaviour, whilst still conveying subtle aspects of its psychological state such as frustration at not being able to get a point across. Having covered emotional states, personality and related social behaviours, section 7 then reviews how a character presents its psychological state through appropriate behaviours including facial expression, gaze, body pose and movement.

To close we distill some key guidelines that summarise the material in the paper. Developers and researchers can use these guidelines to access the relevant sections of this STAR.

2. Why expressive virtual characters?

Virtual humans are an essential part of the content in many types of application such as in entertainment, games and story-telling [Art05, CCM02, MPP01], training environments [GM05, JDR*05, BLB*02], virtual therapy [FSB*03, SPS99, HBR*04, MLL00], conversational representatives (avatars) [GSV*03, MVS*02, VC98], and expressive conversational interactive agents [GB03, RPP*03, CBB*99]. For applications that require only animated not interactive content, there are a variety of tools that can either capture or hand model human behaviour. However, this is very labour intensive and it is only economical when a very specific performance is required, as in, say the movie industry [Las87, Lin01, SS01, Jac03]. When we move to interactive systems, there are simply too many possible situations and responses to make hand-modelling or motion capture feasible. There has to be an element of simulation and modelling.

Once a system is interactive, we must start to model the psychological state of the character, how this is represented, and how this state changes depending on the context and content of the scene including other characters and the user(s). The modelling of affective behaviour and social norms in the virtual human becomes especially important if the application depends on "virtual human-user" interaction involving communication and socialisation within virtual environments [BLB*02, BB04]. This is challenging because, as evaluators or users of these systems, we have specific expectations of how people behave and respond given a situation. People generally expect virtual characters to behave in a manner befitting its appearance and will often be disturb by discrepancies in its behaviour. The interpersonal communication of emotions, interpersonal attitudes, personality traits within individuals is integral to regulating the communicative and behavioural ability of virtual humans [Arg69, Pic97, GRA*02]. It has also been argued that the inclusion of a psychological state and expression may contribute to a richer interaction [Pic97].

In the physical world, interactions with socially unskilled individuals are often incongruent and difficult. Such an uncomfortable encounter can often lead to anxious appraisals by those involved [MD04]. A similar phenomena is observable in virtual environments. An individuals' perceived behaviour realism of virtual characters is positively associated with their experience with it [VBG*04] while a lack of expression in virtual characters has a negative impact on the perceived communication in collaborative virtual environments. This was reflected in statements made by participants in a study which investigated a negotiation task between pairs of participants represented by avatars [VGSS04, GSV*03]:

> *A large part of a normal conversation - especially given the delicate nature of the subject, involves a lot of facial expressions & gestures which play a large part in the conversation... ...After realising the fact that the avatar conveyed very little of the persons actual physical/emotional state, it became even less believable ...*

Evaluative studies suggest that virtual characters can elicit the appropriate, and sometimes surprising, responses from

(a) (b) (c) (d)

Figure 1: *Examples of some highly photo-realistic virtual humans with no or limited interactive ability.*

participants [SVS05]. This is particular true of responses and situations that have a strong emotional content. For instance, individuals respond to praise, criticism, personalities and social responses from a computer or a virtual character in the same manner they would in response to another individual [KMP02, NB03, RN96, SPS99]. Prendinger *et al.* [PMI05] found that a character whose body language expressed empathy with a person, and apologised for problems with a computer significantly decreased both the perceived difficulty of the game and the participants' stress, as measured via physiological responses. Slater *et al.* [SPS99] asked people to give a short talk in front of an audience of virtual characters, and found that participants' rating of their own performance at public speaking was significantly affected by whether they were given positive or negative nonverbal feedback by the audience. This is consistent with Reeves and Nass's theory that people generally treat computers as social actors [RN96].

Although we can observe strong reactions to expressive virtual characters, a survey of literature and observations indicates that different people respond to different levels of expressive cues with difference levels of physiological, psychological and behavioural responses [BH04]. This depends on a number of factors including an individuals' ability to perceive & interpret the psychological state of others, their characteristic traits (personality), their emotional state, their anxiety threshold etc. For instance, there is empirical evidence that individuals who are prone to paranoia in the physical world are more likely to get anxious in response to virtual characters in essentially neutral contexts such as those depicted in figure 2(f) [FSB*03, SPBC04]. However, the design of behavioural models for virtual characters is a complex challenge. There is no dictionary to translate emotions and personality into the appropriate behavioural cues and the repertoire of behaviour will necessarily be quite large. There is evidence that a lack of behavioural range and expressiveness can hinder performance in collaborative virtual environments [SS02, TBS*98].

A final piece of the motivation for studying expressiveness is that in the quest to solicit a realistic response, behaviour may be more important than the visual realism of the character. Although visual realism is extremely important to convey aspects of a characters' status (figure 1), it is what the character does that conveys more information. Bailenson and Blascovich have argued that the visual realism of an avatar is only important in that it allows for the generation of social behaviour [BB04], and that the importance of the avatars' behaviour realism far outweighs visual fidelity in some applications [BLB*02]. Further, individuals in a study by Nowak and Biocca [NB03] reported a higher sense of copresence while interacting with a less humanoid representation.

One consensus that is emerging is that the virtual characters' visual realism and behavioural realism need to be balanced [GSV*03, Sch02, TBS*98, VSS05]. Empirical studies conducted on the impact of avatar visual and behavioural fidelity have confirmed this to a certain extent. For instance, the simulation of inferred [GSV*03, GSPR05, Sch02] or expressive [FSB*03, PSB01, SPS99] behaviours in an agent can greatly affect an individuals' experience in virtual environments.

3. Nonverbal communication

Generally face-to-face communication channels can be divided into two distinct but interrelated categories: *verbal* and *nonverbal*. The verbal communication of an individuals' psychological state is undertaken using both literal ('*I am irked, angry or outraged*') and figurative ('*blowing a gasket*') statements. Every verbal message contain two elements, the content and an insight into the relationship between the individuals in the conversation [Duc98, WBJ68]. Nonverbal behavioural changes give a tone to a face-to-face communication, accent it and sometime even overrides the verbal part of the communication [AT80]. Studies have shown that if an unfriendly message is delivered with a smiling facial expression, the message is taken to be friendly [AT80]. Even though verbal and nonverbal content might not always indicate the same message, what they convey is almost always compatible [GRA*02]. However, nonverbal communication is generally taken to be indicative of the true psychological state of an individual especially when the cues are negative [AT80].

If we want to create truly socially believable characters,

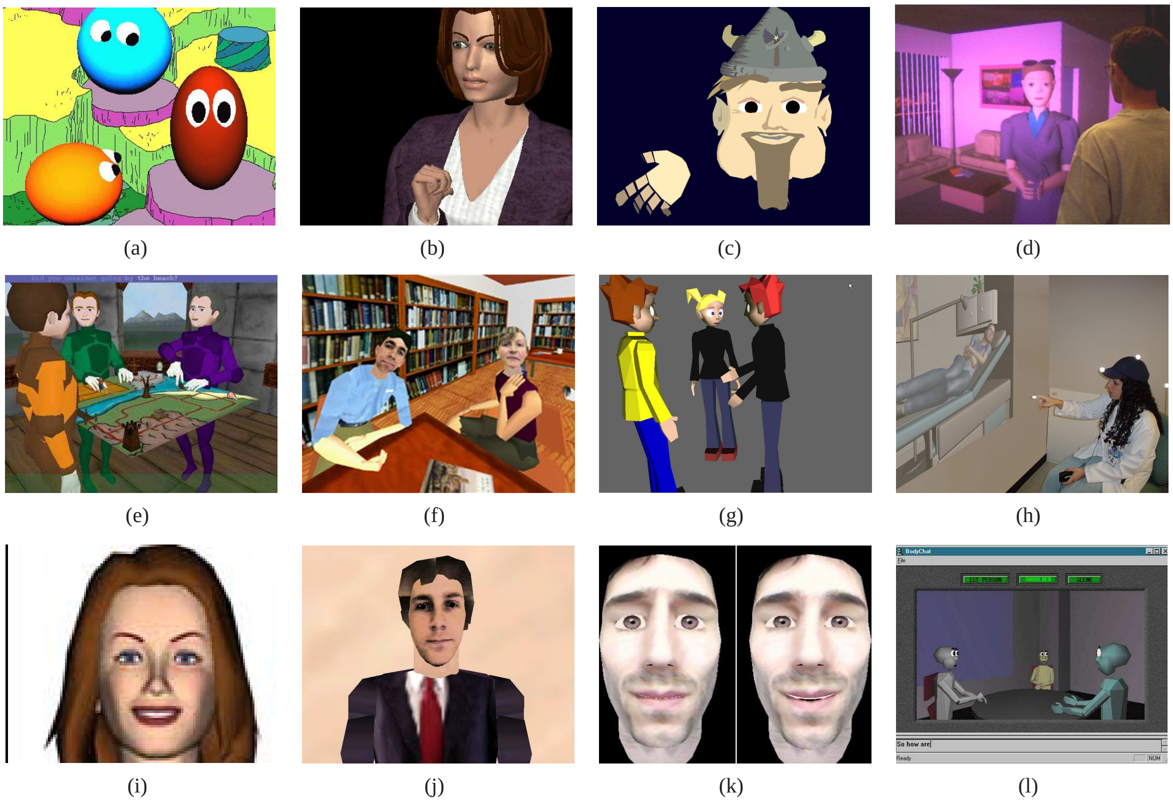

Figure 2: *Examples of some virtual humans with models of nonverbal behaviour.*

nonverbal communication is a vital element to bring into it. However, nonverbal behaviour is highly complex and it is important to understand all of the issues before trying to build it into a character. Even simple behaviours like the *smile* depends on many factors including *culture*, *interpersonal relationship* and *context*. Behaviour can have different meanings in different contexts, for example looking some one directly in the eye can be loving or highly aggressive depending on your relationship. Knowledge of the cause and context within which a nonverbal behaviour is expressed can greatly expand the interpretation of the behavioural cue [AT80,Ekm65,Sch64]. This means that it is important to take into account the context in which a character will be used in. Culture is also a very important factor [DRPP04] and it is important to be sensitive to cultural differences when deploying characters internationally.

Nonverbal communication can also depend on whether someone is conscious of and intends to send a signal. For example, people can pretend to express an emotion they do not feel, a fake smile, or unintentionally express an emotion they are trying to hide, like an interview candidate whose words sound confident but whose posture reveals nervousness. Wiener *et al.* make a distinction between *signs* and *communication* [WDRG72]. Communication is a goal directed action while signs are behavioural responses. Distinguishing between signs and communication has a lot to do with the awareness of both the individual expressing the message and the other. While communication is always directed at another person a sign can either be due to another person (e.g. involuntary laughter) or something that is unconnected with other people (making a hunched posture when cold). The difference between these three cases can be thought of in terms of how a character could react to a player in a game. If a character is unaware of the players presence it will just produce signs which are undirected to the player whereas if it is aware of the player it will use communication. However, if it is trying to ignore the player it will not communicate but some involuntary signs aimed at the player might be "leaked".

Nonverbal communication also has many different functions that it is important to understand when trying to create a character. The functions can be of different types. The first type is the expression of a persons' mental state, for example:

- Provision of information about the individuals' feelings or attitudes to others in an interaction [MF69]. Nonverbal behaviours function as cues in the expression and intensity of the emotion or affect [PPL86]. These are often called *Affect Displays*.

- Nonverbal behavioural cues are often used to either project a personality trait such as dominance or unconsciously provide cues to the individuals' persona [Meh71, MF69]
- The information can also be about someone's cognitive state. For example, looking away or expressing concentrating can indicate reflection, while other facial expressions can show understanding or confusion [PP02b, PP02a]
- *Adaptors* are self-touching behaviours that provide insight into an individuals' attitude and anxiety level since they are the least consciously monitored behaviours. They are mainly used unconsciously for instance wringing your wrists and therefore provide a rich source of involuntary information about the individual.

Of this type the first two, emotion and personality are possibly the most important and are discussed in the next two section.

As well as being about a persons' internal mental state nonverbal expressions can also be about relations between people, as described in section 6. We divide this type of behaviour into two, firstly behaviour that conveys information within a conversation and secondly behaviour linked to inter-relationships. Typically the second type of non-verbal behaviour as identified through conversation can be described by the following cases:

- The regulation and effective management of a conversation. Individuals engage in a complex sequence of head nods, glances and nonverbal vocalisations closely synchronised with verbal communications [Arg98, PPL86]. These behaviours are called regulators.
- Emphasising certain syntactic and semantic features of speech [CPB*94]
- *Emblems* are standardised gestures and signals that are well understood in a particular culture. They often used intentionally and consciously in situations when verbal communication is not possible for instance waving to indicate departure from a noisy scene.
- *Illustrators* on the other hand are signals that are invented on the spur of the moment but that are still voluntary, have a clear meaning in addition to speech. An example might be using a gesture to show the curve of a road. Illustrators often have an adjectival function, describing in more detail what is being talked about. A special class of illustrators are deitics, pointing gestures (though some deitics are involuntary and so are not illustrators).
- Nonverbal communication can also have a *performative* function, showing what action a statement is trying to achieve, as opposed to its literal meaning. For example the performative of "I think you should phone John" could be advice, a request or an order.

The final type of behaviour gives information about long term relationships between people, as further described in section 6.2:

- Expression of intimacy and emotional closeness. Close proximity, touch, gaze, reduced verbal fluency, longer pauses between utterances and increased silence distinguish romantic relationships from friendships [Duc98]. Other observations in social psychology [Fus02] and anthropology
- Self presentation and social control including attempting to project an attitude of dominance or agreeableness. Individuals find it difficult to disagree with others sitting beside them [Duc98, PPL86].

It is very important to understand the various functions and how they interact when implementing an expressive character. For instance, Thórisson and Cassell [TC99] conducted a study comparing a character (Gandalf, figure 2(c)) that was only capable of conversational nonverbal communication to one that was only capable of emotional expression. They found that people communicated better with the character with conversational behaviour. This shows that it is important not to focus solely on one expressive feature such as emotion without considering certain communicative functions of nonverbal behaviour. However, this result should not be taken to mean that emotional expression is irrelevant. The study involved talking about an emotionally neutral subject (astronomy). Garau *et al.* [GSBS01, GSV*03] performed studies that featured characters with conversational behaviour but no emotional expression. They found that when engaged in emotionally heated negotiation, participants found the lack of emotional expression problematic. A possible conclusion from these studies is that some basic conversational behaviour is fundamental for characters but that many other types of behaviour, such as emotional, are also important. Another conclusion is that what forms of expression are needed is highly dependent on the context in which a character is used. For instance, emotions and relationship might not be important in a banking setting, but both are vital for a character used for a long term health care counselling application.

As well as different function nonverbal communication has many different *modalities* of expression: facial expression, the eyes, body posture and movements. Most functions can be expressed in different way through different modalities. Once we have discussed the various functions of nonverbal communication, we will describe the different modalities in section 7.

4. Emotions

Everyone knows what an emotion is; until asked to give a definition [FR84].

Emotions are loosely regarded as a reaction to personally significant events where the reaction may include biological/physiological arousal, changes in cognitive processes, behavioural/social/motor expression, action tendencies and subjective labelling of these feelings [KK81]. Disney and other cartoonists have maintained that perceived emotional states are a necessary substrate for producing plausible characters [TJ81]. The character of Grumpy in Disney's version of "Snow White and the Seven Dwarfs" would not be the same without his regular expression of irritation and anger. Creating emotion is essential to creating perceived intelligence and reasoning in agents [Dam95, Min88, Pic97]. Scheutz suggests a number of potential roles for emotions and moods in agents including action selection, goal management, social regulation, learning, and memory control [Sch04]. Emotions guide actions while moods function to shift not only cognitive content but also the individuals' processing mode [ED94, Ise87]. For instance, individuals in a good mood are significantly more successful at solving problems [IDN87] and emotionally charged events are generally more memorable than unemotional events, even more so for negative events [TF85]. The central idea is that emotions and mood are always involved while thinking and should be simulated in virtual characters in order to express plausible behaviour. The addition of an emotional dimension to a virtual human has a significant effect on the interaction, however, modelling them is not straightforward. In addition, to the confusion amongst theorists on the definition and functionality of emotions [KK81], another problem is the lack of agreement on what effects emotions have on behaviours. Any given category of emotions can motivate a variety of expressions and actions. The following sections attempt to summarise relevant theories and categorise existing computational models of emotions.

4.1. The role of emotions

Emotion theorists have taken the view that emotions are responsible for generating a rapid and efficient response to important environmental stimuli which is useful for survival from an evolutionary point of view [GM05]. In general, the primary function of emotions are to guide actions and provide information through facial, vocal and bodily expressions.

On a biological level, emotions prepare the body for actions like the flight or fight response in the face of oncoming threat [Fri88]. In situations where hesitation could have dire consequences, emotions function to set aside cumbersome cognitive processing allowing for strategic planning [LN02]. Emotions create the optimal physiological milieu to support the necessary behaviour in an emotionally charged event. On a cognitive level, emotions alter an individuals' priorities thereby serving to allocate limited resources towards multiple plans and goals [OJL87]. Emotions can arise out a set of a deliberate planning process which in turn can influence the decision making process. The affective system allows for rapid response, efficient social communication and is generally adaptive but is also prone to errors. Once emotional situations escalate, guiding focus to the immediate and relevant goal makes individuals loose perspective thereby leading to irrationality. This is the main argument against incorporating emotional models into agents. It has been suggested that individuals develop coping strategies to manage their emotional states leading to models simulating this relationship in virtual humans [MG02, GM05, PDS*04].

Emotions aid in the communication of an individuals' needs to the self [Fri88] and to others [Lev94]. The intensity of the emotional communication and expression often depends on the status of the interpersonal relationship between the individuals. Individuals often prefer to discuss emotional topics with friends than strangers [BH04] and a group leader who is perceived to be adept in emotional expressiveness is likely to influence others emotionally [For90]. For instance, persuasion is more successful if the individual doing the persuasion is emotionally (*motivated*) about their point. This is a useful aspect when try to simulate a virtual human with variable decision-making abilities such as in circumstances involving stressful situations as described by Gratch and Marsella [GM05] or a virtual human which is trying to influence the users' actions.

An emotional experience is more memorable and involved than a non-emotional one [TI65]. Individuals in a depressed mood have increased accessibility to sad memories and decreased accessibility to happy memories [TF79]. El-Nasr *et al.* use a learning component to define the expectations of an agent using information on past events [ENYI00]. For instance, the generation of a fear response in expectation of an undesirable upcoming event. Lim *et al.* use long term emotional memory and emotional tagging to influence the re-experiencing of events [LAJ05]. Emotional states can be seen as an important factor in retrieving specific memories and also as a useful method to index perceived memories in virtual humans.

4.2. Emotions and moods

Mood represents the overall view of an individuals' internal state. Other than functionality, an affective state is differentiated as an emotion or a mood based on three other criteria: *temporal*, *expression* and *cause*. Emotions are brief lasting for a matter of seconds or at most minutes. Emotions are often associated with a facial expression [EF76] and have identifiable cause. Moods last for longer, are not associated with a specific expression or cause. The main functional difference between emotions and moods is that, emotions modulate actions while moods modulate cognition. In accordance with this definition, modulating the actions caused by an emotion becomes difficult if it occurs during a mood causing cognitive instability. Emotions are phasic changes superimposed on moods which in turn can be thought of as the affective background. Despite these differences, emotions and moods are inextricably linked. Emotions can lead to particular moods, and moods can alter the probability that a particular emotion will be triggered and its intensity [NSS01, ED94]. For instance, an individual in an irritable mood becomes angry more readily than usual and the resulting anger is more intense, decays more slowly and is more difficult to control. There is no research to determine if this is because the individual is in a continually low level of anger and readily provoked or because there is a difference in thresholds and related cognitive appraisals characterising the mood. Most existing systems represent moods as a low level of emotional arousal for a longer duration than emotions [Vel97, KMT02a, RPP*03]. Becker *et al.* [BKW04] developed a presentation agent (Max) that expressed a coherent portrayal of emotions over time including the dynamics of emotions and moods over time. Recently Tanguy *et al.* [**?**] implemented a concept of mood along two dimensions (tension and energy) as suggested by Thayer [Tha96].

4.3. Intensity of emotions and emotional decay

The intensity of emotions is affected in accordance to a set of variables which include how important the event is, level of unexpectedness, prevailing mood, gender and arousal [Fri88, GW93]. An increase in either of these factors intensifies the emotion. An increase in the period of time from the eliciting event results in emotional decay. Emotional decay is explainable by the view that emotions are elicited not so much by the presence of external stimuli but by an actual change in the stimuli [Fri88]. The intensity of an emotional state in a virtual human is generally implemented in some form in existing computational model of emotions [Vel97, Gra00, ENYI00, PI01]. Many computational models use emotional intensities as one of the ways to create emergent personalities in the agents [PDS*04] while others use a model of personality in conjunction emotions to create agent emotional states with differing intensities to the same event [PI01, EKMT02, Geb05]. For instance, in Bates *et al.*'s Woggles [BLR94] (figure 2(a)) and Rosis *et al.*'s Greta [RPP*03] (figure 2(b)), emotional intensity is assigned depending on the uncertainty in the agents' beliefs and the importance of achieving a goal.

Generally the relationship between physical expression and emotional intensity is modelled in a linear fashion. This is in keeping with results reported by Hess *et al.* which indicated that the perceived intensity of the underlying emotion of morphed natural faces is linearly related to its actual physical intensity [HBK97]. In addition, they found that the more intense the emotional expression, the more accurately the target emotion was recognised by the individuals in the study. This result goes some way in explaining the success of exaggerating expressions to enhance the plausibility of virtual characters [PVP*03, TJ81]. On the other hand, Bartneck and Reichenbach [BR05] recently found a curve-linear relationship between the physical intensity of synthetic faces and perceived intensity. They also report similar results to Hess *et al.* in that the recognition accuracy of an emotional expression increases with physical intensity but *only* up to a certain point beyond which the accuracy levels do not vary significantly. This suggests that exaggeration of behaviours will only be worthwhile to a specific point.

4.4. Participant responses to emotions

Moods and emotions of others around an individual influences their emotional state and subsequent behaviours. There is also evidence that behavioural feedback influences the emotional state of individuals and postural feedback may intensify the emotional experience [DLS*89, FJLC99, HCR94]. This process of *catching* the prevailing mood of others is known as emotional contagion [Doh97]. Individuals have a tendency to automatically mimic and synchronise facial expressions, postures, and movements with those of another person [GK97], which leads both individuals to converge emotionally [HCR94]. A group of people who are good at perceiving emotional expressions are more likely to tune into the emotions of those around thereby resulting in a sense of togetherness [BH04]. The presence of either positive or negative emotions in an individual prompts others to act in a manner suitable to most likely lead to a state of equilibrium in which both individuals reach a neutral mood set [Fri88]. This could be a useful property to simulate in a situation involving groups of virtual humans.

Studies suggest that the emotional state of a virtual human can similarly affect individuals in a virtual environment. Mortensen *et al.* [MVS*02] reported that participants in physically remote locations, who were collaborating with a confederate in a shared virtual environment, were able to assess the mood state of the confederate despite the simple avatar. The recognition of the confederates' mood was largely based on tone of voice, but also on the disposition of the avatars' body, for instance, a drooping head indicated depression [MVS*02]. Emotional contagion was also observable in a study conducted by Pertaub *et al.* in participants who were asked to give a speech to a virtual audience [PSB01, PSB02]. Participants who gave a speech to the negative audience developed greater anxiety. Participants also judged the negative audience has being more realistic than the positive audience [PSB02]. This has significant implications in the design of virtual humans employed for therapeutic applications [FSB*03].

Another aspect of emotional contagion is to do with the increase in affiliation between individuals sharing same situation. Gump and Kulik [GK97] reported that in keeping with Schachter's '*emotional similarity hypothesis*', threat increased affiliation and did so when the participants believed to be facing the same situation. Schachter [Sch59] argued that physiological arousal evoked by a threat causes uncertainty, and therefore individuals in similar threatening situations desire to affiliate with others to evaluate the intensity, nature,and appropriateness of their emotional state. In an empirical study involving the virtual recreation of a bully-victim type situation, young participants felt empathy towards the victim and felt anger towards the bully especially when the participant identified with the victim [PDS*04, PDS*05, HWA*05]. Hall *et al.* reported that expressions of empathy were increased when young participants had high levels of interest in the conversation between the virtual characters [HWA*05]. In other words, the more young participants perceived that a conversation was believable, the more the felt sorry for the character. Further results indicated that if the young participants perceived that they were similar to the virtual characters (identified with them), they expressed greater empathy and liking [HW05]. The concept of the emotional similarity hypothesis might prove to be especially relevant under collaborative scenarios of a particularly stressful nature such as those explored by Gratch and Marsella [GM04] or in situations calling for strategic collaboration in games.

Exposure to an emotional episode colour the perception of an individuals' view of the world along numerous other dimensions. Happy individuals make more positive judgements whereas sad individuals make more negative judgements. It is one of the most reliable affective phenomena [FB87]. Prendinger *et al.* reported that an empathetic virtual human had a positive effect on the participants' perception of the difficulty of the task and significantly decreased stress caused through the delay [PMI05]. The ability to project emotional expressiveness gives the impression of a more trustworthy, charismatic and credible individual and the impression of a charisma. This is aided by others paying more attention to behavioural cues in order to obtain feedback on the progress of the situation [LN02]. This property could be exploited in a useful sense when trying to design virtual agents in an e-commerce setting where trust and credibility play an important role.

4.5. Models of emotions

The categorisation of emotions is as fuzzy as the definition of emotions. For instance, there is little doubt that anger and sadness are emotions but there is less agreement on moods (irritability, depression), long-term states (love), dispositions (benevolence), motivational feelings (hunger), cognitive feelings (confusion, *deja vu*) and calm states (satisfaction) [BH04]. Gratch and Marsella [GM05, GRA*02] categorise approaches to emotion modelling into two main groups: communicative-driven approaches and simulation-based approaches. Instead of modelling an internal emotional state, the communicative-driven systems focus on picking an appropriate display of perceived emotions based on end-goal of the agent. A lot of communicative-driven models use a model of basic emotions such as those defined by Ekman [EF78]: happiness, surprise, disgust, fear, sadness and anger. This approach is well-suited for applications aimed to acts as educational aids or e-commerce representative where the emotional range of the virtual agent is well-defined and relatively narrow. The simulation-based systems attempt to simulate an internal state (emotion) which is used to choose the appropriate expression. Even though the end result is often aimed to be communicative, this approach affords more flexibility and variety in the expression. Most simulation-based systems are based on an appraisal theory of emotions such as the "Ortony, Clore and Collins" (OCC) model [OCC88]. This approach allows the consequences of events on a variety of levels to be appraised in accordance to the goals, standards and attitudes of the agent before resulting in an emotional state. It works well for applications requiring decision-planning in a group of virtual humans (agents) or in circumstances requiring a human-virtual human interaction over a prolonged period of time.

4.5.1. Models based on basic emotions

The concept of basic (or pure) emotions was made famous by Ekman and Friesen [EF78] and is commonly applied using morph targets to simulate emotional expressions [HBK97]. Each model proposing a case for basic emotions have its own set of basic emotions [OT90]. The six basic emotions as defined by Ekman [Ekm82] were associated with a set of facial expressions [EF76].

Ekman's complete set has been used in a number of early systems such as Velásquez's Cathexis [Vel97] implemented on a 2D baby face - Simón [Vel98]. Each basic emotion had a family of related affective states in order to implement emotional intensity; for instance, fear was associated with fright, terror and panic. Cathexis also allowed for the modelling of emotion blending; for instance, grief was a mixture of sadness and anger or fear depending on context [Vel97]. Currently the Ekman's set of basic emotions (or a sub-set) are utilised as a method of categorising the end-effect of expressing emotional states after the internal emotion is appraised through more complex systems. These include Kshirsagar and Thalmann [KMT02b] and André *et al.*'s models of emotion and personality [AKG*99], Ushida *et al.*'s emotional model based on fuzzy inference rules [UHN98], Rosis *et al.*'s Greta [RPP*03, RPP*03] (figure 2(b)), and Tanguy et al's [?] Dynamic Emotion Representation (DER) model (figure 2(k)).

The most noticeable restriction in Ekman's set is the imbalance between categorising negative and positive affect [Ekm82]. In order to overcome the unbalanced nature of Ekman's set of basic emotions, El-Nasr *et al.* [ENYI00] created a set of emotions for evaluations of their Fuzzy Logic Adaptive Model of Emotions (FLAME) on a synthetic baby face, which included sad, anger, joy, fear, laughter, pain, thirst and tired. Another restriction in the Ekman's set (as well as other basic emotion models [OT90]), is the lack of sufficient labels to represent a rich set of emotional expression. Rosis *et al.* [RPP*03] get over this by using more than two models of emotions: the Ekman's set [Ekm82], a sub-set of emotion labels from the OCC model [OCC88], and embarrassment and shame [KB96].

Schachter and Singer [SS62] reported that participants were more susceptible to the mood of a confederate when they had no other explanation for an increased psychological state of arousal. This suggests that the context of a situation plays an active role in the emotional state felt by individuals. This is in keeping with results reported by Ushida *et al.* [UHN98] where participants reported more perceived emotional states than the actual six that were represented in very simple agents. Several personalities and motivations were attributed to the agents including the basic survival type motivations (thirst, appetite, and feeling good) and more complex social motivations (defending territory and communication). Freeman *et al.* [FSB*03] reported similar results in which participants attribute sentience and had feelings of paranoia towards completely scripted virtual characters (figure 2(f)). This implies that for some applications simple emotion models might suffice in producing plausible behaviour. In addition, to Ekman's model of basic emotions, there are a number of other models [OT90] such as Plutchik's model which allows for contains four pairs of *opposites*: joy and sadness, acceptance and disgust, fear and anger, surprise and anticipation [Plu80]. Plutchik's theory is more balanced than Ekman's set of basic emotions, allows for emotional blends and varying emotional intensities (rage is more intense than anger). Albrecht *et al.* [ASHS05] uses an emotion model, based on the "emotional wheel" described by Plutchik [Plu80]. In this model, the emotional space is represented by a disk defined by two dimensions: activation and evaluation. Similarity between two emotions is proportional to the angle that separates their positions on the wheel. The emotional wheel model is used also by other facial animation systems, including [KSS96], [RKK03], and [LAAB02].

4.5.2. Models based on appraisal theories

Another common approach to modelling emotions is to view the emotions as reaction which result from appraisals/assesments of events and objects in correspondence to goals (and probabilities of achieving them), beliefs, risks and attitudes. Appraisals can be basic sensory-information processing, can involve rapid and automatic cognitive processes or a much slower cognitive process. Plutchik's model goes some way in forming such a chain [Plu80] but a more commonly used and comprehensive appraisal model is the OCC model [OCC88]. The OCC model provides a rule based system for triggering 22 emotions, however, this has been judged to be too complex for modelling in a virtual human [Bar02]. Ortony [Ort01] revised the emotional structure of the OCC model to 10 containers by eliminating all the branches relating to the concerns of other virtual humans. The argument is that the slight reduction in realism is a justified tradeoff for some applications given that the revised emotional structure reduces the computational complexity of the model.

Bates, Loyall and Reilly [BLR94, Rei97] built one of the first emotional agent (Woggles, figure 2(a)) system on an architecture called Tok which consisted of Hap (action selection) and Em (emotional model). Em generated emotions based on the success/failure of goals as appraised by Hap. Em was built based on the OCC model and emotion intensities were generated based on the importance level of the goal. For instance, hope and fear in agents are the result of the belief that a goal had the chance to succeed or fail. In addition, to generating emotions, Em also represented basic interpersonal relationships between the agents on a like-dislike dimension. For instance, the proximity of a disliked agent to an agent causes it to become angry [BLR94]. Ushida *et al.* [UHN98] also present an emotion model for simple spherical agents with a deliberative system based on the OCC but use a set of fuzzy inference roles which control the levels of seven emotional factors. Gratch [Gra00] built on the work done on Reilly's Em algorithm and further extended it to produce a generalised plan-based model of emotional reasoning in Émile. Émile allowed agents (in this case a pedagogical agent - Jack and Steve [RL98]) to appraise the emotional significance of events in relation to its own goals and the probability of achieving those goals. Émile was integrated in Marsella *et al.*'s IPD [MLL00] (Interactive Pedagogical Drama) system which, amongst other things, focused on the impact of emotional states (and intensities) on virtual human behaviour [GM01, MG01]. Another way in which the significance of goals and beliefs of achieving those goals is represented by Rosis *et al.* [RPP*03, RPP*03]. Greta was equipped with a representation of beliefs and goals (based on a BDI architecture [RG91]) that drove the generation of emotions and the decision to display the expressions. The internal states of the agent were generated through the use of Dynamic Belief Networks. Greta's internal states allow for changes in emotional intensity with time, response delays, blends [RPP*03].

El-Nasr *et al.* [ENYI00] suggested an approach to modelling the dynamic nature of emotions by simulating their effects on behaviour by using a learning process to activate blends of emotion that would affect and be affected by a number of factors including motivation. FLAME is based on an event-appraisal model which uses fuzzy rules set to map assessments of the impact of a certain event on pre-defined goals into an emotional intensity and state [ENYI00]. The generation of emotions were again defined using an event-appraisal model based on the OCC. The agent learns about the properties of different events through reinforcement learning and about the user through a probabilistic model that keeps track of the users' actions. Evaluations of the agent (PETEEI [ENIY99a]) resulted in some emergent behaviours, for instance, an oscillation/confusion between extreme emotional states [ENIY99b]. However, the learning model significantly improved the plausibility of the affective behaviours expressed by the agents [ENIY99a]. Kesteren *et al.* [KAPN00] follow the same principles as El-Nasr *et al.* and simulate natural emotional expression through the modelling of the OCC model using neural networks.

OCC-based appraisal models have been used in conjunction with other mechanism like coping [GM04] or social networks [PI01]. Gratch and Marsella [GM04] focused on intensely stressful scenarios and therefore extended their unified model with the addition of a detailed model of problem-focused and emotion-focused coping to produce EMA (EMotion and Adaptation). In addition, a simple personality model is used to allow the agent to choose which coping strategy it prefers to deal with a particular stressful situation. The coupling of the appraisal model with coping process models has led to some unexpected emergent coping behaviour [MG02]. Similarly Paiva *et al.* [PDS*04] created a 2D cartoon-like autonomous humanoid character application (FearNot!) with an emotional model coupled with coping mechanism to evoke empathy in participants. Young participants, especially females, felt increased empathy if they perceived the virtual characters to have followed their coping strategy advice [HWA*05]. Prendinger and Ishizuka [PI01] build their work on the premise that agent behaviour can not be generated by modelling internal states such as personalities, attitudes and emotions alone but has to integrate social role awareness models. This allows their agent to suppress the expression of an emotional state if it would result in the failure of a goal.

Few constructive additions have been made to the OCC model. Bartneck [Bar02] argued that a function of history should be incorporated into the model so that the same event occurring again would not result in the same emotional intensity. However, this is dealt with in the models using values to represent the desirability of specific goals/events. Picard [Pic97] and Bartneck [Bar02] point out that the OCC model is not designed for and therefore does not allow for interactions and dynamics between the different emotional states. Models based on the concept of basic emotions or the OCC model support emotional expression. The interactions and dynamics between emotional states is covered in models described in the next section which deal with mechanisms which elicit emotions as well.

4.5.3. Models based on primary, secondary and tertiary emotions

Damasio [Dam95] and Sloman [Slo01a, Slo01b] categorise emotions into primary, secondary and tertiary emotions. The definition of primary emotions is similar to the definition of basic emotions in that they are defined as being innate. Primary emotions are produced by reactive mechanisms mapping external stimulus patterns to behaviours. For instance, the states that often elicit two major response patterns, '*fight or flight*', are anger or fear respectively. Secondary emotions, such as hope, are learnt associations between recognised stimulus patterns generated primary emotions and analysed situations where these patterns occurred [Dam95]. Tertiary emotions arise from the interaction of emotions and other cognitive processes (e.g. motivation) or through limited resources while pursuing multiple goals.

Scheutz *et al.* [SSL00] introduced the CogAff agent architecture which models agents' cognitive system into a reactive, deliberative (what-if processes) and meta-management (reflective process) layer. Primary emotions were triggered in the reactive layer, secondary emotions were triggered in the deliberative layer and tertiary emotions involve the meta-management layer. Evaluations suggested that in a simulated survival-type scenario, agents with reactive mechanisms and affective states could achieve the same goals more efficiently than agents with high-level deliberative mechanisms [SSL00]. More recently Tanguy *et al.* [?] presented the Dynamic Emotion Representation (DER) model which represented changes over time in emotion intensities and the interactions between different emotions (figure 2(k)). Emotional *impulses* produced by the mechanisms eliciting emotions, such as those based on the OCC model [OCC88], effect (and are effected by) the state of the DER. Primary emotions are used to trigger pre-organised behaviours that are associated to facial expressions as defined by Ekman [Ekm82]. Secondary emotions based on the Ekman set to select facial signals corresponding to communicative functions. For instance, an individual with a high level of happiness might emphasise a word by raising his eyebrows where a person with a high intensity of anger might frown to achieve the same result. Figure 2(k) shows two types of smiles generated by the system depending on whether the character is sad or happy. Tertiary emotions, represented using the mood model of Thayer [Tha96], are used as filters on how emotional impulses effect primary emotions and how they change the intensities of secondary emotions.

The DER model is built over models which elicit emotions from internal or external events such as those defined earlier. Models based on Lazarus's proposed process involving primary appraisals, secondary appraisals and reappraisals [Laz91] allows for a much more dynamic representation of emotion process.

5. Personality

> *The fact that Bugs Bunny says "What's up doc?" in his distinctive Brooklyn accent is part of his personality [Rei97]*

Moffat [Mof97] differentiates between emotions and personalities over two dimensions: duration and focus. Where as emotions are temporally inconsistent, personalities remain constant and is not specific to particular events. Personalities arise out of more indirect and long-term factors. An emotion is a brief, focused change in personality.

Personality represents the unique characteristics of an individual. In psychology, the aim is to represent and understand human psyche. This is done through defining various dimensions to generalise possible personality traits amongst individuals and scale them in some way. Many of these models have been used to create personality in agents

[BB98,KMT02a,Geb05]. This approach to personality modelling helps in designing virtual characters that have certain characteristics which partly determine their behaviour. For instance, a friendly virtual character is likely to act friendly in any situation because of the traits in its personality. Isbister and Nass [IN00] reported that participants were able to identify the personality of agents and preferred an overall consistency in agents regardless of whether the agent was matched to the personality of the individual or not. This suggests that agents with a friendly smile are expected to maintain friendly and open body postures. Granted this detracts away from modelling personality quirks as strong as Bugs Bunny [Rei97], although it does given a starting point and the task of creating quirks is left to an artist.

Personality traits come into play when a virtual human is used in an application that is meant to create some sort of relationship with an individual or in cases where a group of virtual humans are placed in social setting.

5.1. Models of personality

A number of models focusing mainly on emotions tackle personalities by modelling emergent personalities. For instance, Ushida *et al.* [UHN98] model various personality types through the difference in emotional expression. For instance, the threshold levels for triggering an angry state in the agent is used to control the extent to which an agent is irritable. In Rosis *et al.*'s Greta (figure 2(b)), personalities were implemented by varying the goal weights that change the importance agents attach to each goal. However, applications, which involve running a virtual human over a substantial period of time, call for a more robust personality model [RPP*03].

Generally personality traits are used to set the threshold to generate emotional states and control the intensities. Information about the characters' personality can influence the probability of choosing actions explicitly [PG96] or with algorithms which introduce uncertainty [CS04, KMT02a]. Chittaro and Serra [CS04] present a goal-oriented approach to modelling agents. The personalities of the virtual humans are modelled through a probabilistic automata (Probabilistic Finite State Machines - PFSM) where behaviour sequences are chosen from an animation library (and sometimes modified) based on personality. Most systems simulating the internal states of agents include detailed models of both emotions and personality since the two are closely linked [KMT02a, AKG*99]. Two of the most prevalent personality models used in modelling individual characteristics are the five-factor model (FFM) [MJ92] and the PAD dimensional model [Meh80].

The five factors that make up the FFM are Openness to experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. The FFM is sometimes referred to as the OCEAN model. Individuals with high *openness* are prone to have more interests with less depth in each interests. Openness to experience is important to creativity. A highly *conscientious* individual focuses on less goals and exhibits self-discipline. Individuals can be classed as focused, balanced or flexible. High *extraversion* refers to individuals who are comfortable with a higher number of social relationships. Individuals can either be extroverts, ambiverts or introverts. Extroverts talk first, talk more and are more persuasive [AT80]. An individual with a high *agreeableness* factor is prone to be subordinate and accept the groups' norms with ease and is termed an adapter. A challenger with low agreeableness factor is more interested in their personal goals. The negotiator is an individual with moderate agreeableness. Finally *Neuroticism* is the negative emotionality factor. An individual who scores low on neuroticism requires more stimuli of higher strength to feel negative emotions and is termed resilient. An individual with a high score is reactive. Neuroticism is also associated with high levels of anxiety [AT80]. The relationship between personality and affective states is not emphasised in the FFM. This explains the coupling of the FFM with the OCC model in many existing systems.

Chittaro and Serra [CS04] use the FFM of personality as input to a probabilistic automata based behaviour animation system [CS04]. Breese and Ball [BB98] modelled agents' level and intensity of happiness with two personality traits (dominance and friendliness) in a Bayesian Belief Network (BBN) model. More recently Kshirsagar and Thalmann [KMT02b, KMT02a] used BBN to model personality traits using the more well-rounded FFM coupled with a layer of mood. They argue that the model handles abstract concepts within a structured probabilistic framework and also handles uncertainty with respect to the generation of emotion. Personality was represented along a n-dimensional space (FFM) while emotions were represented as levels of arousal through an extended version of the OCC. Kshirsagar and Thalmann added two other emotions (surprise and disgust) to the existing OCC framework. The mood of the agent was controlled through a probabilistic function of the agents' personality. The overall emotional state of the agent depended on the emotional impulse caused by an event, the personality, the mood, time-linear emotional decay, and the pervious level of emotional arousal of the agent [EKMT02]. Egges *et al.* [EKMT02] extended this model and linked it to a dialogue system (modelled using Finite State Machines) represented by a 3D face (figure 2(i)). This model is further detailed as the PME model [EKMT04] and was also integrated with an idle motion synthesiser to create idle motions appropriate to the emotional state of a virtual human [EMT05].

Similarly André *et al.* [AKG*99] presented an integrated model of emotions based on the OCC [OCC88] personality based on the FFM [MJ92]. The integrated model was used as filters to constrain a decision process to control an agents' behaviour. Like Rosis *et al.*'s Greta, André *et al.*'s agent was also built on a BDI-based [RG91] architecture. Initially their model simulated agent behaviour based on 4 basic emotions (anger, fear, happy and sad) and 2 dimensions of personality which relate to social interactions (extraversion and agreeableness). Prendinger and Ishizuka considered the same two dimensions of personality [PI01]. André et al then developed their model to include two affective information processing channels: reactive and deliberative. This is similar to the first two layers of Scheutz *et al.*'s CogAff architecture [SSL00]. The deliberative channel generates secondary emotions in accordance to the OCC appraisal model [AKG*99]. Gebhard [Geb05] presented ALMA which focuses on the temporal variations in affective states. ALMA (a layered model of affect) simulates short, medium and long term affective states through modelled emotions, moods and personality respectively [Geb05] based on the OCC model and FFM. Like Kshirsagar and Thalmann's model, the personality traits of the agent is used to control the computation of the intensity levels of emotional states. Romano *et al.* [RSH*05] model social knowledge in addition to modelling personality (FFM) and emotions (OCC). The main disadvantage of using the OCC model and FFM in conjunction is that there is no clear mapping between the two.

Mehrabians' three-dimensional Pleasure-Arousal-Dominance (PAD) model allows modelers to input some of the links between personality and emotions [Meh96b, Meh80]. Different emotions and personalities are viewed as a variations along these dimensions. For instance, a score of low pleasure, high arousal and high dominance would be interpreted as anger while a score of low pleasure, high arousal but low dominance would be interpreted as fear. Some emotion modelers have chosen to reduce the dimensions in the PAD model to just two: pleasure and arousal, following Russell's "circumplex" model of the facial affect space [RFD97] instead of Ekman's Facial Action Coding System (FACS) model [EF75], however, there are studies that argue that two dimensions are insufficient to completely handle aspects of facial affect [SERS00]. Becker *et al.* [BKW04] focus on the modelling of a coherent portrayal of emotions over time in an agent - Max. The emotional engine behind Max consists of two components. One to simulate the dynamics of emotions and moods over time. The other component acts as a emotion categorisation model based on the PAD model [Meh96b].

Instead of using Mehrabian's PAD model to simulate personality, Gebhard's modelled moods in ALMA [Geb05]. Then Mehrabian's mapping between the PAD model and the FFM model [Meh96a] was used to define the agents' personality. Gebhard suggest future additions to the ALMA model to include changes to the agents' mood in accordance to the emotional intensity. Alternatively there are other dimensional models of personality. For instance, Lim *et al.* [LAJ05] modelled personalities of agents along the three dimensions of Eysenck and Eysenck's model [EE85]: extraversion, neuroticism and deliberativeness.

6. Interpersonal factors

The previous two section have discussed the expression of factors that can be considered internal to an individual, their personality and emotional state. However, much nonverbal behaviour is closely related to the interaction and relationships between individuals. Nonverbal communication, is of course, a form of communication and is an integral part of inter-personal communication. In order to build believable characters that are capable of social interaction it is vital to endow them with realistic nonverbal communication and the ability to uphold nonverbal social norms.

There is evidence that the social norms, especially in terms of spatial behaviour, observable in the physical environment hold in the virtual environments as well [MB99]. In fact, some researchers argue that the perception of agent realism is improvable just by upholding social norms in a virtual environment [MB99]. In social situations, purely rational virtual humans prove to be insufficient since the focus is not on providing the best solution to a well-defined problem but rather to produce an suitable output within context. Slater, Tromp, Steed and colleagues observed this in their studies on small group behaviour in a shared virtual environment [SS02, TBS*98]. Even with the simple avatars, individuals were hesitant to seem rude to other avatars by breaking social rules. If behaviours that broke the social norms were observed, such as an invasion of interpersonal space, it was described as uncomfortable and disconcerting [Sch02]. Guye-Vuillème *et al.* conducted a study in which participants had avatars with sophisticated nonverbal behaviour. They found that people made great use of this behaviour to uphold real world norms and relationships. This same effect is overwhelmingly pointed out in the studies conducted by Slater *et al.* on collaboration in a shared virtual environment [SS02, TBS*98]. Individuals in the study specifically mentioned that the lack of avatar expressiveness hindered their joint performance. Casanueva and Blake carried out a study with groups of three individuals in a shared virtual environment on desktop and concluded that avatars with gestures and facial expressions yielded higher reports of copresence than static avatars [CB01].

We divide the use of nonverbal communication for social behaviour into two broad classes. The first class is connected with the details of social interaction and conversation. This is the form of nonverbal communication that takes place in every social interaction, normally over relatively short time periods. The second deals with creating and expressing social relationships between people, a longer term process.

6.1. Social Interaction and Conversation

The most important form of human communication is, of course, spoken conversation, it provides the centre piece of all our social interactions. Nonverbal communication is a vital part of face-to-face conversation, providing an further channel of communication beyond the purely linguistic one.

Cassell, Pelachaud *et al.* [CPB*94] identify three basic classes of function of nonverbal communication in conversation:

- Syntactic functions: functions that depend on grammatical aspects of speech.
- Semantic functions: functions that emphasise, complement or add to the meaning of speech
- Interactional functions: functions that regulate the flow of speech.

Syntactic functions are perhaps the simplest of the three. Nonverbal communication is used to emphasise grammatically important elements in speech, for example raising eyebrows or gesturing at an accented syllable in a word or at a pause. Conversely semantic functions are probably the most complex, particularly in their interaction with language, as they reflect the great diversity of meaning that can be communicated in conversation. The final function of nonverbal behaviour is its interactional function, the regulation of the interaction between the participants of a conversation, which will be discussed in the next section.

6.1.1. Conversational Structure and Turn Taking

The purpose of the interactional functions is to ensure that the participants are able to engage smoothly in conversation. As such, probably the most fundamental function is to determine whether, and when, two individuals should engage in conversation at all, in the starting, or initiation phase of conversation. When people approach each other they firstly use gaze, brief glances followed by eye contact to indicate willingness to talk. Once this initial contact has been made they exchange greetings, either verbal ("hello") or nonverbal (smiling, waving or shaking hands) [Ken90]. If some one wants to end a conversation they indicate this by looking around, shifting their attention away from the conversation. Once the other participant acknowledges this they exchange verbal and nonverbal farewells. Vilhjálmsson and Cassel have implemented these conversation initiation and ending behaviour in their BodyChat system [VC98] (figure 2(l)). Peters' characters [PPB*05] are able to reason about each others' attention while approaching and during conversation and use this information to know whether to start of end a conversation

At a finer timescale, nonverbal behaviour also regulates behaviour within a conversation. The major interactional function is turn taking, the process which determines who should speak at any given time. At any given time each participant in a conversation has one of two roles, the speaker or the listener. These two roles can simply be modelled with a finite state machine, such as Colburn *et al.*'s [CCD00] simple system or Thorisson's sophisticated model [Thó02]. In conversation people successfully take turns at speaking, with only rare overlap, and often as little as 100ms between one person stopping speaking and the next starting [Thó02]. This brief time interval alone indicates that there must be some sophisticated method at work to regulate the transition between roles.

The transition can happen in two ways. The speaker can give the floor to the listener at the end of their utterance or the listener can interrupt the speaker, and if the speaker acknowledges this interruption the listener has the floor. In the first case the speaker must indicate that they wish to give the floor to the listener, which is done using a number signals such as a long gaze at the listener at the end of an utterance or lowering pitch.

If the listener wishes to take the floor from the speaker, they must indicate that they are ready to start speaking. This can simply be done by starting to speak, an interruption. However, nonverbal signals can make the process smoother and less jarring by warning the speaker before the listener starts to speak for example a sharp intake of breath or beginning to gesture are both associated with starting to speak. Gesture can also be used by the speaker, dropping hands to a resting position can signal the end of a turn while keeping them in the gesturing position during a pause can keep the turn.

Cassell and colleagues have developed a number of virtual characters that exhibit realistic turn taking behaviour [VC98, CBB*99, Vil05]. Thórisson has also a developed a system that recognises a system that is able to recognise turn taking signals from a real real person, enabling smooth conversation between real and virtual humans [Thó97, Thó02].

6.1.2. Speaker Behaviour

The nonverbal behaviour of speakers is intimately connected with their speech and so when implementing a virtual character it is important to be aware of how speech and behaviour relate to each other.

Much of the role of speakers' nonverbal is to express a diverse range of semantic information. One of the most important distinctions, made by Poggi and Pelachaud [PP00] based on Speech Act theory, is between the propositional and performative meaning of a communicative act. The propositional meaning is the literal meaning of the spoke words, generally this is contained in the verbal channel, though sometimes not entirely, for example deitic function (pointing with hands or eyes) might be part of the propositional meaning. The propositional meaning can be about the outside world (pointing or descriptive gestures) or about the speakers mind (beliefs, emotions of thoughts). On the other hand the performative value of the act is the action that the speaker intends to achieve. For example, if one were to say to a colleague: "I could really do with a coffee", the performative might be what is indicated by the literal, propositional value of the sentence: informing her of my desire of a coffee. However, it is likely to have a different performative, for example, a request that she comes to the common room and join in drinking a coffee, and perhaps an offer to buy her a coffee.

Nonverbal communication plays a strong role in displaying the performative value of a statement, for example, distinguishing a piece of advice and an order. The performative meanings have many components including the goal (giving information vs a request); whose goal it is (the speakers or listeners), the degree of certainty and the power relationship between speaker and listener [PP00]. For example a request in the interest of the speaker with high speaker power and certainty is an order while with low speaker power and certainty and in the interest of the listener can be advice.

There are three ways in which meaning conveyed in the nonverbal channel relates to that conveyed in the verbal channel (and for that matter, how the meanings of different modalities of nonverbal behaviour relate to each other) [PP00]:

- Repetition and redundancy: the verbal and nonverbal channels give the same information.
- Substitution: a nonverbal element takes the place of a verbal element, for example, a pointing gesture can take the place of a full description of a place, or a smile can take the place of a greeting.
- Contradiction: the verbal and nonverbal meanings are opposite. This happens in irony and jokes. It can also be the result of nonverbal "leakage" of suppressed affect or suppressed meaning in a lie [RA05a].

The problem with generating semantic nonverbal communication is that the range of possible signals is very large, and the meaning of each signal is very specific. An implementation of such behaviour would require a lexicon of signals that would approach the size and complexity of the lexicons used in natural language. However, we lack the standard dictionaries and corpora that are available for verbal language, and it may not even be possible to create such dictionaries as it seems that gestures can be invented on the spur of the moment.

Another, and simpler to implement, use of nonverbal communication in speech is to mark emphasis. A number of modalities can do this, perhaps the most obvious being vocal intonation, with vocal stress appearing on emphasised words. Gestures are also very important in marking stress, beat gestures (see section 7.4 below) have been found to occur simultaneously with vocal stress [Arg98]. The eyes can be used to mark emphasis, particularly the "eyebrow flash" [PP02b]. One of the most successful models of emphasis for virtual characters has been Cassell and colleagues' use of discourse structure [CPB*94]. They divide utterances into a *theme* part and a *rheme* part. The theme is roughly what the utterance is about, the part of the utterance that refers to subject matter that is already known in the conversation. The rheme is the new information in the utterance. The main emphasis of the sentence is normally on an important word in the rheme and Cassell *et al.* [CPB*94] propose generating gestures simultaneously with verb phrases in the rheme. They have since extended this work to posture [CNB*01] and gaze [TCP97].

6.1.3. Listener Behaviour: The Back Channel

The nonverbal behaviour of the listener is as important as that of the speaker, particularly as it provides feedback to the speaker. While the most obvious flow of information in a conversation is from the speaker to the listener there is a significant flow of (largely nonverbal) information from listener to speaker that is called the *backchannel*. Via the backchannel speakers gain important information about whether the listener is paying attention, has understood, and how they have reacted to what is said.

The importance of listener behaviour is demonstrated in a study by Slater *et al.* [SPS99]. Participants were asked to give five minute presentations to a virtual audience of eight male agents dressed in formal wear and of varying behaviours: a positively charged set, a mixed response set, and a negatively inclined set. The positive demonstrated high levels of attention, gazing at the speaker 90% of the time, leaned forward and smiled frequently. On the other hand the negative audience avoided eye contact, slumped in their chairs, yawned and fell asleep. The mixed audience started with a negative attitude and ended with a positive attitude. Participants were asked to rate their own performance at speaking. This rating was correlated with the good mood of the audience. This indicates that the backchannel behaviour of listeners can have a strong effect on the speaker.

Kendon [Ken90] suggests 3 types of back channel signal:

- "listening behaviour" which indicate attention and understanding (or confusion)
- "feedback behaviours" which give information to the speaker about the listeners reaction to what is said (for example, whether they agree).
- "imitation", listeners often imitate the speakers' behaviour in some way.

Of these various functions attention is generally shown through gaze. Understanding and confusion can be shown through facial expression, gaze, nonverbal vocalisations (e.g "uh-huh") or even full words (saying "yes" without interrupting the speaker). This type of backchannel facilitates achieving mutual understanding of what is said and what is meant, the common ground, the process of *grounding*. Nakano *et al.* [NRSC03] have implemented a virtual character that models grounding in conversation with a real person. During an explanation it looks for signals of understanding in the persons speech or gaze, and only continues to the next point when these have been found

Another interesting aspect of backchannel behaviour is imitation or *interactional synchrony* [Ken70]. This is the tendency of a listeners' movements to become synchronised with those of the speaker. In the course of an interaction the beginning and end of the listeners' movements, for example shifts of posture, will tend to occur at the same time as those of the speaker, even if the movements themselves are very different. The movements of both speaker and listener tend to be synchronised with speech, the boundaries of the movement coinciding with the boundaries of words and phrases. This last observation tends to suggest that the synchronisation process is primarily driven by speech.

There have been many virtual character systems that demonstrate back channel behaviour. In order to achieve back channel behaviour the character must be able to detect the behaviour of the speaker, for example using positional trackers and microphones if the speaker is a real person [MGM05, GS05]. A microphone can be used to detect pitch and loudness of speech that can trigger a head nod [MGM05]. Interactional synchrony can also be simulated, the boundaries of an utterance can be detected using a microphone [GS05] and postures shifts using a head tracker [MGM05, GS05], both events trigger a posture shift in the character. Vilhjálmsson and Cassell [Vil05] use a different approach. In their system users interact with each other via animated avatars. The users type in text which is spoken by the avatars. The text is parsed for information relevant to nonverbal communication. This information is used to animate the speaking avatar but is also sent to all listening avatars to generate appropriate backchannel behaviour.

6.1.4. Systems

This section will discuss how these conversational factors can be combined together into complete architectures for characters that are capable of conversation, often called Embodied Conversation Agents (ECA). An embodied conversational agent must be capable of both verbal and nonverbal behaviour and so should have use a representation that combines both. Markup languages provide a powerful method of embedding nonverbal information in (verbal) text. The Avatar Markup Language(AML) [KMTGV*02] directly embeds nonverbal behaviours, such as facial expressions into the text. The Affective Presentation Markup Language (APML) [DCCP02], on the other hand takes the approach of adding communicative functions, such as performatives or turn taking behaviour, in the markup. This allows the text to be combined with other markup information specifying information about the character and context and use this to generate the final behaviour [RPPN02, MLP04].

As discussed above a conversation character can either be an agent, whose behaviour is entirely computer controlled, or an avatar, that represents a real person and whose behaviour is partially controlled by that person. An agent can engage in conversation with other agents (figures 2(a)) or with real people (see figure 2(d), 2(e), 2(h) and 2(l)). We will not discuss agents that only interact with other agents [GB04, JT05] can avoid many of the issues relating to real time recognition of a real persons' behaviour and so are a subset of those that interact with real people. Our example

Figure 3: *Cassell* et al.*'s architecture for Embodied Conversational Agents*

will be an agent architecture proposed by Cassell and colleagues [CBC*99, CBB*99, CBVY00]. The architecture is based on their FMBT model:

- F: separation of propositional and interactional Functions
- M: Multi-modal input and output
- B: separation of function and behaviour (a function might be "greet" while the behaviour might be say "hello" or wave)
- T: real-Time interaction

As shown in figure 3, the architecture must first use inputs from the conversational partner to understand their behaviour, then choose a response and finally generate appropriate behaviour. The input manager handles the input sensors which consist of audio to detect onset of speech and pauses, a speech recognition system to determine what is said and a computer vision system to detect gestures. An understanding module takes these inputs and interprets them based on the current discourse context and a static knowledge base. The reaction planner generates high-level discourse functions that should be generated and the generation module transforms these into concrete behaviours. These three modules are all deliberative and so may not be fast enough to produce the highly time dependent behaviour needed in conversation so they are controlled by a reaction module that is able to produce faster responses without the need for the deliberative modules and manages various information such as the turn taking state. Finally, an action scheduler takes the outputs of the reaction module and generation module to produce synchronised speech and body language in an agent such as Rea (figure 2(d)).

Avatars generally have their speech provided by the person controlling them. In order to generate appropriate nonverbal behaviour they must either parse this speech [CVB01, Vil05] or take further user input [GB04, GVCP*99, VC98]. As an example of this type of system we take three iterations of Vilhjálmsson, Cassell and colleagues' system for conversational avatars. BodyChat (figure 2(l)), an online chat system where people were represented by avatars with realistic conversational behaviour [VC98]. This used minimal user input, conversational text and information about whether to start and end a conversation and generated gaze and gesture behaviour. This behaviour included conversation initiation and ending, turn taking and backchannel behaviour. BodyChat was a successful system, in a user trial [CV99] users found an avatar with this autonomous behaviour more natural and expressive than one where the control was manual. More interestingly people that used the avatar with autonomous behaviour actually felt more in control than those that had (complete) manual control over their avatar. However, nonverbal communication is closely related to what is said, and BodyChat did not use any information from the text itself. This lead problem Cassell, Vilhjálmsson, and Bickmore [CVB01] developed BEAT, a system that parses natural language in order to generate nonverbal behaviour. BEAT parses input text and tags it with grammatical information and dialog structure. It then uses these tags to suggest appropriate nonverbal behaviour, using a number of rules. It initially suggests all possible behaviours and then uses a set of filters (which could depend on the personality and culture of the speaker) to choose which are actually used. BEAT was used to create a conversational avatar system, Spark (figure 2(e)), in which users type text, which is then parsed by BEAT to generated complex nonverbal behaviour [Vil05]. Spark was shown to be able to predict over 50% of a persons' nonverbal behaviour from what they say. In user trials it significantly improved quality of conversation and performance in a collaborative task.

6.2. Social relationships

If people need to interact with characters over long time periods it is vital that we create characters that are able to form social relationship with humans, what Bickmore and Cassell [BC01] call *relational agents*. This is particularly important in areas where trust is important such as sales [CB03] or where self disclosure is required, for example health care [BP04].

One important aspect of relationship is the attitude of one person to another. Argyle [Arg98] proposes a model based on two dimensions: friendly vs. hostile (affiliation) and dominance vs submission (status). Behaviours associated with high affiliation include physical closeness, more direct orientation, more gaze smiling, head nods and lively gestures. Low affiliation (hostility) is harder to define as there are normally strong social inhibitions against showing hostility in public, however, closed postures such as crossed arms, and lack of smiling can show hostility. Bécheiraz and Thalmann [BT96] have a model of posture generation based on affiliation. In general high status can be characterised by height and size, people draw themselves up to their full height and use large postures, expanding the chest and having hands on hips. Low status individuals tend to be more hunched over and use smaller postures. In established hierarchies, however, a relaxed posture can indicate dominance. Prendinger and Ishizuka [PI02] have developed a model of the evolution of attitude over time. Their characters' attitudes to each other depend on their current actions but also the history of their past actions, and whether the two are consistent.

Another aspect of Argyles model is the equilibrium theory: people seek a certain degree of social closeness (linked to affiliation), which is the result of a balance between a tendency towards attraction and one towards repulsion. If people become too close on one modality they will tend to compensate, if two people are placed very close together they will tend to look at each other less and vice versa. Gillies and Ballin [GB03, GB04, GCB04] implement Argyle's model, including both affiliation, status and the equilibrium theory applied to posture and gaze (figure 2(g)).

Bickmore and Cassell [BC05] have a more complex model of social closeness consisting of three dimensions:

- *familiarity*, the degree to which people have exchanged information about each other. As relationship progresses people exchange more personal information.
- *solidarity*, how similar two people are to each other in their interests, beliefs and behaviour.
- *affect*, the degree of liking, analogous to Argyle's affiliation.

Their character can use various strategies to increase each component. Though most of these strategies are verbal they use some nonverbal cues such as interactional synchrony, smiling and nodding.

As well as relationships between individuals, people can have relationships with whole groups of people. Prada and Piava [PP05] model individual status and affiliation relations between characters but each character also has a position within a group (how they are considered within the group) and a motivation for that group (how the group considers them).

6.2.1. Trust

Trust is another important dimension of attitude. For many applications of virtual characters it is vital that people trust the character, for example if it is selling a product or giving medical advice. Nass *et al.* studied the effects of agent personality in conjunction with ethnicity and found that individuals attributed more trust to an agent perceived as an extrovert (more expressive) with an identical ethnicity [NIL01]. There has also been work done in agents with respect to building trust and a social-emotional relation with individuals by remembering past interactions with the participants and anticipating their future needs [BC01]. There has also been some recent methodological work done by Cowell and Stanney in exploring the particular behavioural cues and emotional expressions that aid the enhancement of credibility in virtual humans [CS05]. These studies suggest that incorporating positive and persuasive behaviours in virtual humans increases their perceived credibility and the individuals' propensity to trust in them. Rehm and André [RA05a] take the opposite approach, creating a character that shows nonverbal behaviour associated with lying. Examples include *masking*, putting on a false expression such as a smile. These false expressions tend to have different (longer) timings and are more asymmetric than normal expressions. Though their results were mixed there is some evidence that these behaviours reduced trust.

6.2.2. Politeness and Social Inhibition

Most of this paper discusses how people express their feelings. However, in many social situations we cannot express our feeling directly as it might be considered impolite. This is particularly true of negative emotions and attitudes. Prendinger and Ishizuka [PI01] have created a model of social inhibition. They define a measure of *social threat* which is a combination of *power* (equivalent to Argyle's dominance) and *social distance*, a measure of how well two people know each other. When talking to some one with a high social threat the intensity of a displayed negative emotion is reduced. The intensity of display is also reduced by the characters' agreeableness personality attribute (see section 5).

Politeness is also an important factor in social restraint. In Brown and Levinson's model [BL78] politeness is about avoiding threats to other peoples' *face* (the self images that people would like to project in public). Everyone has two types of face the *positive face* is the want to appear desirable to others and the *negative face* is the want to act independently and have ones' actions unimpeded by others. A threat can be a criticism (positive) or a request (negative) and threats can be avoided by appealing to the positive face (e.g. a compliment), negative face (a less direct request "would you mind opening the door"), or by an off the record statement ("I am tired" meaning "stop the meeting"). Politeness strategies have mostly been applied to virtual characters in terms of verbal behaviour [BC05, GRW05], but Rehm and André [RA05b] have found some associations between the use of gestures and politeness strategies.

7. Modalities of Expressive Behaviour

There are two main issues involved in studying the expressions of internal states. The first revolves around the question: is there a set of distinct expressive behaviours for an emotional state caused by a specific stimuli? The second issue concerns the challenge of identifying commonalities in the expressive behaviours used with a specific emotional state caused by *different* stimuli. The same emotion is expressed differently by different individuals at different times [Arg69], however, individuals are able to use even minimalistic cues to perceive the emotional state of others. In other words, are there distinct basic cues that can be used to perceive an emotional state in others?

Individuals express conversational feedback, various emotional states and interpersonal attitudes with a range of behaviours including: facial expression, movement, gestures, head orientation and eye gaze, posture, vocal intonation, and linguistic expressions. Generally these modalities are used in conjunction and are highly synchronised. For instance, Planalp conducted a study to ascertain the variety of cues used to perceive emotions in naturally occurring situations [PDR96]. Majority of the participants (97%) reported using more than a single cue. On average six to seven cues were used, thirteen cues being the maximum number reported. Two thirds of the individuals reported using vocal cues while over a half used facial, indirect verbal and context cues. In terms of accuracy, 84% of the participants correctly perceived a single emotion and 68% matched all emotions in cases where multiple emotions were felt. The following sections give an overview of various nonverbal behaviours used in the physical world, their functions and how existing virtual human behaviours have been tackled.

7.1. Facial Expression

One of the most expressive areas of the body is the face capable of producing about twenty thousand different facial expressions [Bir71]. It is the area most closely observed during an interaction [Arg69]. Facial expressions of an emotional state is readily recognisable by others even in a synthetic static sketch format [WKSS00, HBK97]. Facial expressions have been well studied and categorised by researchers in accordance to main expressions [Ekm82, EF78, WKSS00], responses [Osg66], and basic physical movements [Bir68]. Most methodological research has focused on the importance of facial expression [EF75] or accompanying physiological reactions [Can15] of emotions at the expense of body movement (gestures), body orientation (turning away), body agitation (tenseness), and speed of movement. The display of emotions through the face is only one side of the coin, facial expressions are also used to support the speech, adding new or redundant information [BC97, PB03, PP00]. The following sections deal with the types of facial expressions and their significance.

7.1.1. Describing Facial Expressions

It is important to understand how facial motor primitives and facial signals combine to create facial meanings in order to understand what facial expressions communicate. Smith and Scott [SS97] summaries three approaches to describing how facial expressions communicate meanings: the *purely categorical model*, the *componential model* and the *purely componential model*.

The *purely categorical model* describes a limited set of full-face patterns often associated to universally recognised emotional facial expressions. Ekman [Ekm92], and Izard [Iza71] conducted studies across different cultures. They showed pictures of full-face configurations and asked participants to associate each picture to an emotional label. From these studies, a limited set of universally recognised facial expressions were identified such as Ekman's six basic emotions [Ekm82]. In accordance with the purely categorical model, meanings can only be expressed through a full-face configuration. Even if the facial patterns could be described by its component using FACS, each component does not have any meaning by itself. Two main problems exist with the purely categorical approach: the first one is the limited set of possible meanings associated with each facial configuration, creating an imbalance in comparison with the recognised communicative power of human faces [SS97]. The second problem is that full-face configurations do not occur often under normal circumstances [Rus97]. These problems are dealt with in the componential-based models. In the *componential model* suggested by Smith and Scott [SS97], meaningful units are facial component actions which are related to Action Units (AU) described by FACS (see section 7.1.3 for a discussion on FACS and AUs). In the *purely componential model* the meaning of the full facial pattern is equal to the sum of the meanings of its components. In contrast, the *componential model* suggests that the meaning of the whole might be different than the meaning of the sum of its parts [SS97].

7.1.2. Functions of Facial Expressions

The face is the most expressive medium of communication in the human body [Arg98, Arg69]. Two main theories explain the functions of facial expressions, the first one sees them as a *reflection of mental states*, such as emotions, and the second one considers facial patterns as *Communicative Acts*.

For a very long time facial expressions have been seen as a reflection of mental states and the most common association is between facial patterns and emotions. Collier [Col85], Ekman & Friesen [EF69] and Smith & Scott [SS97] argue that people look at the face to find cues to emotional states. Emotions and emotional intensities are perceivable from an individual's face with a great deal of accuracy [EF78, HBK97]. Darwin [Dar72], Tomkins [Tom80], Ekman [Ekm92], and Izard [Iza71] argue that emotions, particularly basic emotions, produce typical facial patterns. Ekman [Ekm92], and Izard [Iza71] suggest that these sets of facial expressions are recognised across cultures as expressions of emotions. However, an experiment by Tomoko *et al.* [Tom04] shows that even in-between Chinese and Japanese cultures, some expressions can be interpreted with different meaning. For instance, the expression 'surprised' for Japanese is interpreted as 'Confused' by Chinese.

These facial displays occur during emotionally significant events. They are uncontrolled by the individual and have particular physical and timing patterns related to the emotions they communicate. The resulting pattern is often considered a full-face configuration and are seen as such due to the method used to study them [Ekm92, Iza71]. Izard [Iza97], Ekman [Ekm99] and Smith & Scott [SS97] define a loose link between facial displays and emotions: certain facial patterns are shown during certain emotional states but the presence of emotions is not necessary for these facial patterns to be shown. Also the presence of emotions is not sufficient to produce the related facial displays. It is worth noticing that facial expressions co-occurring with emotions are not exactly the same as those displayed when the emotion is not present. There is a difference between a fake and genuine smile [Ekm92].

Carroll and Russell [CR96] argue that the context within which a facial expression is shown determines the emotions being perceived from the face in question. If individuals are shown a static emotionally expressive face with no clue to what elicited it, they imagine a context [PK02b] and if the nonverbal behaviour is contradictory to the context, individuals will try to justify it [AT80]. Therefore the context within which the emotional facial expression is displayed seems to play a significant role on how it is perceived.

Facial expressions emerge from the communicative process function as part of *communicative acts*. A communicative act is composed of two parts: a communicative meaning/function and a signal which is the physical realisation of the meaning [BC00]. For instance, facial signals such as *eyebrows raised* relate to *emphasise* (see Table 1).

Table 1: *Functions of conversational facial action [BC97, Page 342, Table 15.1]*

Meaningful Displays by speakers			
	Semantic		
		Redundant	
			Personal reaction
			Portrayal
			Thinking
		Nonredundant	
			Personal reaction
			Portrayal
	Syntactic		
		Grammatical markers	
			Emphasiser
			Question marker

Communicative acts are synchronised with speech adding new and/or redundant information. Communicative acts can be acted, symbolic or intentional, and arise due to the communicative process in contrast to emotional expressions which arise due to emotional events. Communicative acts are also fast due to their synchronisation with the speech, in comparison to emotional expressions which have their own time signatures. The meanings of communicative acts are varied. They could be syntactic, semantic [BC97] or related to the person's goals [PB03]. Facial expression, as one of the communicative behaviours, depends on the communicative goal the person wants to pursue [PPdR*05]. The choice of the goal is determined by contingent events and long-lasting features. Contingent events includes content to communicate, felt emotions (physical resources), context and interlocutor. Long-lasting features refers to the agent's culture, personality and style, etc.

7.1.3. Levels of Description for Facial Expressions

Facial expressions in virtual characters can be considered on four different levels: *Static Appearance*, *Basic Motion Primitives*, *Facial Movements/Signals* and *Facial Meanings* from lower level to higher ones in that order.

Static Appearance: This description refers to a 2D/3D mesh, or the volumetric representation of a face. They are not dealt with in this STAR.

Basic Motion primitives: In real faces the basic motion primitives are the movements of the muscles of the face. In graphical characters, motion is controlled by deforming the underlying meshing, either directly [Par72], or by using more sophisticated techniques such as techniques such as Bézier or B-spline patches and abstract muscles [PW96, Bui04]. To obtain the natural movement of human facial expressions using linear interpolation of facial mesh data is not sufficient so other methods are often used. For example, Parke [Par72] used a cosine interpolation scheme to approximate the acceleration and deceleration of facial movements

One of the major contribution into the low-level analysis of facial expressions is the FACS developed by [EF78]. This system is based on Action Units (AU) describing visible movements on the face generated by the contraction of muscles. An example of AU could be "Inner Brow Raiser" or "Lip Corner Depressor". It was developed as a standard method to code facial movements from images or videos but now it is widely used in computer animation. A less detailed implementation [Tho96] provides a simpler way that allows animation with a more cartoon style look. In this approach, control points such as 'Brow Right Medial' and 'Mouth Left' were selected to maximize the expressively/complexity trade-off. The FACS is very useful to represent facial appearances but it does not provide any information about the meanings communicated through facial expressions. Action units have been built into a number of facial animation standards. The most famous ones are the Facial Animation Markup Language (FAML) and the low level Facial Action Parameters defined by the MPEG-4 standard [HEG98, PK02a, GSW*01]. These commands contain intensity information used for positioning a parameter but no information about their intensity changes over time. The exception is the FAML which defines a duration of activation.

Facial Movements or Facial Signals: Commands used for facial movements are based on the same descriptions as the previous one. In addition, functions of time are used to describe intensity changes over time. Generally these movements are defined over three periods of time: attack, sustain and decay duration [DRS02, KMMTT91, Bui04, PB03]. EMOTE and FacEMOTE presented by [BAZB02] are interesting solutions to change the expressiveness of a character (see section 7.3.3 for more details).

Facial Meanings: Implementations of facial expressions at this level are no longer physical descriptions of the face but descriptions of meanings expressed through the face. Example of languages developed to script meanings into facial expressions are Affective Presentation Markup language (AMPL) [PB03], the high level Facial Animation Parameters (FAPs) defined by MPEG-4 standard [HEG98, PK02a], Gesture Markup Language (GML) and Emotion Markup Language (EML) [GSW*01].

7.1.4. Emotional Expressions in Animation System

Numerous facial animation systems display emotional expressions based on variants of a set of universally recognised facial expressions such as Ekman's [Ekm82, Ekm92]. These systems either display one set of the universally recognised facial expressions [KMT02a, Bui04, Vel97, PDS*04] or in addition they produce combinations of these facial expressions [ASHS05, KSS96, RKK03, LAAB02, PM03].

For the selection of emotional expressions, certain systems use a *static emotional representation* [ASHS05, KSS96, RKK03, LAAB02, PM03], such as the emotional wheel described by Plutchik [Plu80]. The emotional wheel describes an emotional space through two dimensions: evaluation and activation. Basic emotions corresponding to the universally recognised facial expressions are mapped onto this space. To display a new emotion, this emotion is mapped onto the wheel using the two dimensions and the combination of the universally recognised expressions corresponding to the two closest basic emotions is shown.

A drawback with static emotional representation is the lack of consistency mechanism. The use of static emotion representations enables a system to display an expression of anger consecutively to an expression of happiness. Emotions vary relatively slowly so opposite emotional expressions cannot occur consecutively. A minimum period of time should separate the two expressions. Certain systems use dynamic emotion representations which are responsible for the production of emotional expressions [Bui04, KMT02a, ?]. These types of systems represent the slow changes of emotional intensities and therefore provides a consistency mechanism to produce emotional expressions.

In the architecture of the agent called Obie, described by The Duy Bui [Bui04], Ekman's basic emotions are used to generate emotional facial expressions. Two fuzzy rule-based systems map an emotional state vector, provided by an emotional model, to facial expressions: one to map a single emotion to its universally recognised facial expression, and the other maps two emotions with the highest intensities to a blend of their universally recognised facial expressions. An "expression mode selection" selects if one or two emotions should be displayed. If the difference between the intensities of the two highest emotions is greater than 0.5 only one emotion is expressed. The blending of the facial expressions is carried out by expressing the two emotions in different parts of the face as defined by a set of rules.

The work described by Kshirsagar and Thalmann [KMT02a] uses of an emotional model to select facial expressions (see section 5.1 for a discussion of the emotional model). As input, the system takes a list of possible responses associated with an emotional state probability provided by a chat-robot system called ALICE. For each emotional state probability, the system computes the probability of each possible mood state in relation to the personality of the character and some thresholds. The mood state with the higher probability is selected and used to choose one of the emotional states and its associated facial expression.

In the Emotionally Expressive Facial Animation System (EE-FAS), emotional expressions are based on the universally recognise facial expressions [?]. Emotional impulses are sent to the Dynamic Emotion Representation (DER), triggering emotional behaviours expressed through emotional facial expressions. The thresholds of emotional behaviour activations vary with the state of the DER providing some consistency to the mechanism producing emotional expressions.

Mancini *et al.* developed a system in which a visual feedback is given to the user using a graphical representation of a human face [MBP05]. Using the real-time extraction and analysis of acoustic cues from the music performance, the system provides an interpretation of the emotional intention of the performer which is then represented as facial expression on an empathic agent. It uses a set of dimensions of expressivity [HMP] to modify the animation of the agent qualitatively.

7.1.5. Communicative Functions in Animation System

Facial expressions are extremely important in interpersonal communication, providing many communicative functions such as emphasis (e.g. raising eyebrows), giving feedback as to whether a listener has understood (e.g. a confused expression) and distinguishing performatives (an order can be accompanied by a frown, a request by a smile) [PP00].

BEAT, described by [CVB01], takes pure text as input to generate "embodied expressive behaviours." The source text is tagged with linguistic and contextual information by the system, which are themselves used to suggest nonverbal behaviours. The system presented by Pelachaud & Bilvi [PB03] and Rosis *et al.* is integrated into a system which translates goals, belief, and emotions into communicative function tags through dynamic belief networks [RPP*03]. During the animation these communicative functions are transformed into facial signals, e.g. physical movements such as a smile.

The first issue due to the use of communicative functions, is that several communicative functions could be expressed at the same time, which means that several facial signals should be displayed at the same time. If these facial signals involve the same facial part and try to communicate different meanings a semantic conflict occurs. The example given by Pelachaud and Bilvi [PB03] presents two communicative functions, a *performative order* and a *comment* which are realised as a *eyebrows frown* and a *eyebrows raised*, respectively. As pointed out by the authors, "adding these two signals would not produce a believable expression", the communicated meaning would be sadness. [RPP*03].

The second issue is the mapping between categories of communicative meaning onto facial signals. "The human face can generated around 50,000 distinct facial expressions, which correspond to about 30 semantic distinctions" [Par02]. As emphasised by Bavelas and Chovil [BC00], several signals could correspond to one communicative function and one signal could be used for several communicative functions. This shows the importance of taking context into consideration when interpreting the meanings of facial expressions. Physical implementations of facial meanings differ according to the physical and mental states of the speaker as well as the state of the dialogue [BC00, PB03].

To solve the mapping problem, the Emotionally Expressive Facial Animation System (EE-FAS) uses a Dynamic Emotion Representation to represent emotional contexts in which physical implementations of facial meaning take place [?]. This technical solution enables the EE-FAS to produce different facial signals corresponding to a category of facial meaning according to the emotional state of the character. Thórisson also solves the mapping issue between signals and communicative meanings by taking into consideration the physical state of the character to select signals [Thó99]. This solution also solve problems regarding semantic conflicts.

7.2. Gaze

The eyes are probably the most intense social signallers in the human face [AT80, LWB00]. Langton *et al.* even suggest that humans have evolved neural mechanisms devoted to gaze processing and gathering relevant information in order to enable the rapid and automated procedures to cope with analysing other individuals' gaze movements and trigger reflexive shifts in their visual attention [LWB00]. These mechanisms are well documented. For instance, during the conversation, it has been observed that individuals listening look at their conversational partner for longer periods of time and more often than the speaker [Arg69, AC76]. Amongst other well known conversation management functions [Kle86], gaze also serves to indicate involvement in an interaction, affiliation towards the other individual in the interaction [EW65], attitude and perceived trustworthiness [Duc98]. Fukayama *et al.* [FSO*01, FOM*02] reported that regardless of visual fidelity, gaze behaviour in virtual characters can be used reliably to convey different impressions to participants.

Individuals prone to seeking affiliation in others, engage in more mutual gaze or glances in search of more involvement in the interaction from the other [AT80]. However, an increase in gaze is also a means of expressing dominance and is perceived as a threat. Exline and Winters conducted studies in a interview scenario in which individuals were placed under either positive, neutral or negative conditions [EW65]. Results indicated that participants in the negative conditions significantly reduced looking at their aggressor (engaging in mutual gaze) [EW65]. In the positive condition, there was a slight increase in mutual gaze while there wasn't any noticeable changes in the neutral conditions [EW65].

Gaze and mutual gaze is also affected in accordance to the emotional state of each individual in the conversation and the interpersonal relationships between them [Ken67]. A concise review of the literature on the functions of gaze and the variations of gaze behaviour in different situations from a social psychology perspective is found in Kleinke's research review paper [Kle86]. The simulation of gaze behaviour in virtual humans is perhaps the most significant part of the face especially in virtual applications requiring social interaction since there is evidence that a fixed stare can cause negative evaluation of the conversational partner [ALC74]. While a simplistic random gaze model might suit the purposes of livening the virtual human, the inclusion of even a simple meaningful gaze model can significantly improve the perceived quality of communication with a virtual human [GSBS01].

A period of mutual gaze between individuals acts as a signal to initiate an interaction causing the individuals to move closer and also to signify attention focus [Gof63, Ken67, MG79]. At the start of the conversation and during the end, the amount of gaze between the individuals is high, however, this levels off to reach a state of equilibrium during the conversation [AC76, AI72]. Kendon observed that the speaker will often avert their gaze when there is a hesitation during the discussion of cognitively difficult material [Ken67] and the rate of blinking reduces [PK27]. Turn-taking is also actively controlled by the gaze and mutual gaze behaviours exhibited by participants in the conversation [TCP97]. In fact Richardson and Dale report results from an experiment using an eye-tracker which suggest that the strength of the relationship between the speakers' and the listeners' eye movements predict the degree to which the listener successfully comprehended the information given by the speaker [RD05]. This implication could be of specific importance to agents used in learning or educational based environments. On the other hand the careful modelling of gaze behaviour in an agent gathering information from an individual could portray the perception of an attentive agent. In a face-to-face interaction, individuals who stare too much can cause others to feel ill at ease or uncomfortable while those who do not look at their conversational partners enough make them feel bored or disapproved of [AT80]. Individuals reduce mutual gaze under negative situations. For instance, mutual gaze is avoided when an individual is embarrassed [Duc98, EW65, AT80]. Again these parameters can be used to induce feelings of threat and dominance in virtual therapeutic applications.

Mutual gaze patterns are also governed through factors other than speech patterns such as seating position, proximity, age and gender [MG79]. For instance, Muirhead and Goldman reported that mutual gaze occurred twice as much when individuals sat opposite each other than beside each other and that middle-aged individuals engaged in half the amount of mutual gaze in comparison to younger or older individuals [MG79]. Individuals also tend to avoid mutual gaze, the closer they are [AC76]. Gaze is also affected according to status and gender; for instance, females dyads of equal status exhibit more mutual gaze than dyads of males or opposite-sex [AT80, MSWB87, EW65]. Females in general engage in more mutual gaze than males. There is evidence that individuals, especially females, tend to engage in increased mutual gaze with the preferred partner (stronger perceived affiliation) in a three-way interaction [EW65].

Deng *et al.* distinguished two approaches taken to model lifelike gaze behaviour models: parametric (discussed below) versus non-parametric or data-driven models [DLN05]. To date there are two main data-driven models. Lee *et al.* [LBB02] observed gaze behaviour in individuals using an eye-tracking device and computed a gaze model based on the first order statistical analysis of the collected data. The model depicted various properties of the rapid motion with which the eye moves from one focused position to another (*saccades*). These properties included the time the eye spends at a focused position (*inter-saccade interval*), the time the eye takes to get to the new position (*saccade duration*), the angle the eye rotates in order to get to the new position (*saccade magnitude*), the direction the eye moves in (*saccade direction*), and the non-uniform velocity it moves with (*saccade velocity*). The resulting model was simulated on a 3D face and participants were asked to judge, amongst other things, how natural, friendly and lively the character was. Overall the model outperformed two control conditions using static and random gaze model. Other than the gaze model, Lee *et al.* also reported a high correlation between the movement of the eyes and eyelids which could, theoretically, be incorporated into an integrated model [LBB02]. More recently Deng *et al.* focus on producing an eye movement model using non-parametric texture synthesis techniques [DLN03, DLN05]. Their technique is based on the observation (also noted by Lee *et al.*) that gaze changes are associated with blinks and considers eye-gaze and aligned eye-blink motion together as an eye-motion-texture sample [DLN05]. These samples are then used to synthesise new eye motions. Evaluations were conducted using the gaze model on a 3D face against the model proposed by Lee *et al.* [LBB02] and a random model. The random model was the least favour model while the model proposed by Deng *et al.* outperformed the model proposed by Lee *et al.* but only slightly [DLN05]. Deng *et al.* suggested that their gaze behaviour model might be well-suited to simulate avatar behaviours to portray boredom or inattentiveness [DLN05]. Rehm and André conducted a study as a probe to obtain data to develop an appropriate gaze model for virtual humans in a multi-way social interaction scenario involving two participants and an agent (Greta) in a game involving deception [RA05c]. There analysis revealed that in general participants followed the gaze patterns observed in dyadic situations and maintained the speaker-listener relationship, however, they gazed significantly more towards Greta when listening to the virtual human [RA05c]. Rehm and André hypothesise that this effect could be due to the difficulties participants faced in interpreting deceptive cues in Greta, the novelty of interacting with Greta or it could be that participants felt more comfortable looking at a virtual human than another participant.

A number of agents have been programmed with parametric gaze behaviour models in conjunction with other communicative gestures (see section 6.1). Colburn *et al.* conducted a study to investigate the differences in an individuals' gaze pattern when interacting with an avatar with varying gaze behaviour models [CCD00]. The gaze behaviour model of the avatar was modelled using state machines triggered in correspondence to who was was speaking and the time passed between states. The gaze behaviour of participants in evaluative studies displayed gaze pattern more similar to those occurring during a real dyad, when there was an avatar with life-mimicking gaze behaviour [CCD00]. Garau *et al.* [GSBS01] conducted a similar study in which, pairs of participants were asked to carry out a negotiation task under four conditions: audio only, avatar with random gaze and head animation, avatar with lifelike gaze behaviour and tracked head movements, and video tunnel. In the two avatar conditions, the individuals in the dyad were represented to each other by identical gender-matched above shoulder avatars. Unsurprisingly, the video was the most favoured condition while the random gaze condition was the least favoured. The most encouraging results were that the lifelike condition was not significantly different to the video condition [GSBS01]. Similarly, Fukayama et al [FSO*01, FOM*02] used a simple gaze model through a two-state Markov model based on inter-saccadic interval, amount of mutual gaze maintained with the participant, and where the virtual character looked when it wasn't maintaining mutual gaze with the participant. Results from their study showed that virtual characters could convey a perceived impression through gaze alone. These studies not only indicate that individuals do respond to virtual humans' with lifelike responses but also that even simple gaze models can elicit these responses. The models employed by both Colburn *et al.*, Garau *et al.* and Fukayama *et al.* were created using the simple guidelines observed by Argyle *et al.* [AC76, AI72, AIAM73] and Kendon [Ken67]. Colburn *et al.* suggested an addition of a transition time-multiplier to their dyadic model to simulate the effects of proximal influences on gaze behaviour when multiple participants are involved in an interaction [CCD00].

Vinayagamoorthy *et al.* created a hybrid gaze model based on the data-driven model by Lee *et al.* [LBB02] and the parametric model used by Garau *et al.* [GSBS01] for use in full-body avatars under immersive settings [VGSS04] (figure 2(j)). The gaze model was evaluated in a study investigating the impact of varying levels of avatar *visual* and *behavioural* realism on participant responses in an immersive virtual environment [GSV*03]. Two levels of behaviour were simulated for the study. The first category of behaviours were designed as a control and included a random gaze model with accompanying arm, legs and head animation. The second category of behaviours included the body animations as well as the hybrid gaze behaviour model [VGSS04]. The higher level of visual realism utilised in the study is depicted in figure 2(j). The study confirmed results implied in the studies run by others [SS02, TBS*98, NB03], that individuals expect more human-like behaviours from more visually realistic avatars. The results of the study concluded that inferred gaze behaviours simulated into avatars representing individuals can have a significantly positive effect. The most interesting interaction effect observed within the factors tested in the study reported that in the case of a lower visually realistic avatar, the inferred gaze model had a *negative* effect on participant response [GSV*03]. In a sense the non-realistic random gaze model was more dynamic and visually stimulating than the subtler eye movements of the realistic behaviour model. Perhaps following in the Disney tradition [TJ81], individuals in a shared virtual environment need to be *made* to believe that a visually simplistic virtual human is "alive" or expressive. This result was especially astounding given that the random gaze model was an averaged out version of the hybrid gaze model [VGSS04].

The above systems model the surface gaze behaviour seen in conversation, but not the underlying processes. Gaze has a primarily perceptual function, directing a persons' attention to an object. Though there has been some work on linking gaze to an underlying model of attention [Cho99, GD02, KVVHJ05], they have not dealt with social interactions. While Gillies and Ballin do argue that gaze can be an important signal to help viewers understand a characters' motivation, this is beyond the scope of this paper. Peters *et al.* [Pet05, PPB*05] are the only researchers that model attention for social interaction. Their character can detect other characters' gaze and from that can reason about their attention and level of interest in a conversation. Based on these two factors the character is able to make decisions about whether to start or finish a conversation.

Poggi and Pelachaud divided the meanings that can be communicated through nonverbal behaviour (see section 6.1) to gaze behaviour and suggest a set of Ekman's Action Units which may correspond to these gaze sets [PP02b]. The geometrical properties tackled by Poggi and Pelachaud is a step towards formalising the animation of gaze behaviour are eyebrows, upper and lower eyelids, wrinkles and eyes. Eye-brows are often used in the expression of emotions like fear, anger, surprise and sad [Ekm79]. These parameters are easily explored using the formalism suggested by Poggi and Pelachaud [PP02b]. Torres *et al.* attempt to further capture the behaviour of gaze with respect to the occurrence of a turn, theme and rheme (see section 6.1 for a discussion on these terms) [TCP97]. Empirical analysis of transcripts of speech, gaze behaviour and head movements were carried out on videotaped dyadic conversations. The results indicated that turn-taking processes are very predictive of gaze behaviour. Torres *et al.*'s results also suggest that the information structure of the conversation accounts for some gaze behaviour [TCP97]. These results were integrated in Gandalf (figure 2(c)) the humanoid agent discussed in section 3. One of the functions defined by Poggi and Pelachaud (deictic) was implemented in an agent called Mack by Nakano *et al.* as part of a model of grounding which included feedback and turn-taking mechanisms [NRSC03]. Preliminary studies suggest that the model encourage more nonverbal feedback when compared to interactions between participants and an agent with no grounding model [NRSC03]. Integrating these parameters into existing gaze behaviour models might prove to be a beneficial research area.

Deng *et al.* argue against the use of parametric model by stating that although the models themselves are compact and economical, they fail to capture important aspects in the data [DLN05]. However, data-driven models are highly customised to specific contexts and involved gathering data over some sort of training period. This calls for a trade-off depending on the type of application. For instance, Deng *et al.* argues against the use of complex analytical models like hidden Markov models to analyse captured data as the hidden states influencing gaze behaviours are not easily interpretable. However, applications based on communication and social interaction could benefit the rich details that could be gained through non-parametric approaches. Evaluative studies have showed that parametric approaches based on careful psychological grounding can have a significant effect on the perceived quality of communication with a virtual human. On the other hand simpler models or customised data-driven approaches may be sufficiently suitable for the design of gaming characters with a short virtual lifespan. To date, there are no complete studies which compare the effects of parametric versus non-parametric models nor are there models detailing the relationship between gaze behaviour and affect in virtual humans.

7.3. The Body

The body can be as expressive a medium of nonverbal expression as the face and eyes. It is particularly important at a distance, Montepare *et al.* [MKZA99] argued that the first cues perceived by an individual when others are approaching to initiate a social interaction are embedded in body movement and gestures. This means that body movement can be an important tool of expression for characters that are to be viewed at a distance. It is also well known that the body plays an active role in portraying personality traits such as dominance or interpersonal attitudes such as affiliation. The same applies to low polygon characters that may not have enough detail for complex facial expressions. Of course, bodily nonverbal communication is also important for high-polygon characters in close up; without plausible body movement, characters can seem stiff.

This section will discuss what can be expressed thought the body and how it is expressed. It is generally accepted that body movements and postures are indicative of an emotional state, however, what is a matter of debate is whether the body is indicative of a specific emotional state (*quality*) or only indicative of the depth of the emotional state (*quantity*). Ekman and Friesen [EF67] suggested that the face conveys specific emotions, while the body conveys the degree of intensity of the emotion [EF67, GBA75]. However, a number of studies and observations have lead researchers to believe that postures can be a more dominant source of cues in the perception of emotions [Arg98, Bul87, DPB65, Jam32, MC03, RGP93] and in some cases an equally accurate source of cues to an emotion as facial expressions [WW86, WW88]. Postural cues are especially thought to play more importance when facial expressions are not viewable, for instance, when the individual is at a distance [WW88]. Walters and Walk report higher recognition accuracy in perceiving happiness, anger and sadness from bodily cues and even more so when movement was involved as well [WW88]. Dittmann *et al.* [DPB65] have reported on the value congruent bodily cues play when accompanying a facial expression of an emotional state. In addition, Dittmann *et al.* also reported that information about the emotional state of individuals was readily available from bodily cues alone.

There are three basic ways in which the body can be expressive. The first is through static posture. The second is through movement. We divide this category into two, firstly how the manner in which normal movements are performed can express emotions, and the second being gestures, a very important expressive medium. The final type of expression in the body is the position in space of the body as a whole,; where we stand relative to other people. In psychology this spatial use of the body is called *proxemics*.

7.3.1. Posture

Charlie Chaplin maintained that if an actor knew his emotion thoroughly he could show it in silhouette [Las87, TJ81]. Quite a number of social psychologists agree with him [Arg98, Bul87, DPB65, Jam32, MC03, RGP93], however, Coulson [Cou04] has shown that some emotions (anger, sadness and happiness) are easier to perceive through the use of posture in virtual characters than others. In addition, the underlying affect of the postures are easier to recognise from the front than from the back or sides. However, in an evaluative study, Vinayagamoorthy *et al.* [VSS06] found that a parametric model of postures based on two sets of Coulson's postures (Angry and Sad) were ineffective in portraying the intended affect in immersive virtual environments. The authors argued that it was better to model full-body virtual humans with casual behavioural cues instead of incorrect or incomplete affective behavioural cues.

Wallbott tested the quality vs. quantity theory on the bodily expression of emotions using two hundred and twenty four video clips recordings of actors and actresses portraying a wide range of emotions of varying intensities. The clips were then *coded* to record the body movements and postures displayed by the actors in the clips using a set coding schemata. Wallbott was able to single out some behavioural cues specific to certain emotions [Wal98]. For instance, an erect posture was rare in cases of portraying shame, sadness or boredom rather the actors used a collapsed body posture. A posture with raised shoulders was typical in cases of anger but moving the shoulder forward was seem to portray fear and disgust. In keeping with earlier studies [Jam32, MC03], the position of the head differed significantly between emotions [Wal98]. The most significant differences were found with the different types of hand and arm movements. For instance, crossing the arms in front were associated with pride and disgust.

It seems that the postures of some parts of the body are more important than other. For example, James [Jam32], found that a forward head and trunk position is taken to mean an emotional expression whereas the outward stretch of the arms and hands suggest a movement and not an emotion. In addition, one part of a total postural configuration is noted at the expense of the remaining posture [Jam32]. For instance, closed/clenched fists indicate tension and the expression of anger. Coulson [Cou04] calls this type of posture an emblematic posture which are likely to be cultural-specific. The head and the trunk of the body were found to be the most significant for the generic expression of the total posture. Secondary factors which were found to be important included the expansion/contraction of the chest and the position of the shoulders (raised/drooped). De Silva *et al.* [DSKBB05] found that Japanese, Sri Lankan and American people agree to a fairly large degree but there were differences in how the intensity of emotion was perceived. Like Coulson, De Silva [DSKBB05] report that some features of postures (often emblematic) had specific meanings to specific cultures.

Kleinsmith *et al.* [KDSBB05] conducted a statistical analysis of emotional posture produced by Japanese actors. Using Multidimensional Scaling they found three main dimensions that explained the variation of posture. They interpreted the first as corresponding to *arousal* that separated sad, depressed and upset from fear, happiness, joy and surprise. Low arousal postures tend to have a bent head and arms placed to the side of the body. The second dimension corresponded to *valence* and separated surprised and fear from happiness and joy. Low valence postures consisted of the head bent forward and the hands raised in front of the face. High valence postures had a raised head and hands held high and away from the body. The final dimension seemed to represent an *action tendency*, with anger being an active emotion while fear and surprise were passive (low action tendency). In passive postures, the hands were raised near the face and the body was kept narrow; whereas in active posture the elbows were bent out to the side of the body and the hands were kept around the hips. While it could be argued that the dimensions Kleinsmith *et al.* found do not correspond exactly to arousal and valence, the model is consistent with other findings. In other studies a lowered head and bent forward trunk was found to correspond to submissiveness and negative emotions such as sadness, shame and humiliation [Dar72, Wal98, MC03, SH95]. On the other hand an upright posture and raised head indicated dominance and positive emotions such as pride and joy [Dar72, Wal98, MC03, SH95]. In studies associated with Kleinsmith, DeSilva *et al.* [DSKBB05] found that posture could also be used to distinguish between different nuances of similar emotions, for instance, the differences between *joy* and *happiness* were particularly notable and consisted mostly in the openness of the arms and distance between the feet. There are exceptions, for instance, Mignault and Chaudhuri mention that in some cases, a lowered head indicated an *illusionary* smile which was interpreted as submission or joy [MC03]. This shows the complexities of involved in modelling nonverbal behaviour. Kleinsmith *et al.*'s model [KDSBB05] is very new and yet to have an impact on virtual characters. It does seem very suited to an implementation especially as there has been very little work on expressing emotion through posture.

As well as emotion, posture can also be important in conversation and representing relationships. In fact posture can be expressive in two ways, the first and most obvious is in the shape of the posture itself while the second is the timing of shifts between postures. Egges *et al.* [EMT05] recorded 10 individuals in conversation as a means of obtaining motion data. In the recorded data, Egges *et al.* reported that most common types of idle behaviour were posture shifts, continuous but smaller postural variations due to breathing or maintaining balance, and supplemental idle motions such as touching of the face [EMT05]. Posture shifts are closely linked to *interactional congruence* [Ken70] (see section 6.1). During a conversation, individuals' posture shifts tend to become synchronised with each other, and with the rhythms of each others' speech thus increasing perceived sense of rapport. This can be simulated in a virtual character by using a head tracker to detect posture shifts and a microphone to detect the start and end of speech in the interactant. This can then be used to program the virtual character to respond with a posture shift [MGM05, GS05]. The synchronisation of posture is likely to be due to a deeper synchronisation of each persons' posture with the rhythm of speech. Cassell *et al.* [CNB*01] have studied how posture shifts relate to discourse structure. They found that posture shifts tend to happen most at the start of discourse segments (change in topic of conversation) and during turn taking (see section 6.1.1). They used these results to implement the behaviour in a character - Rea (figure 2(d)).

As mentioned, the shape of a posture can be important during social interaction, particularly in showing relationships. This type of postural modelling in multi-party conversation has not been applied to virtual characters yet. Scheflen [Sch64] identifies three dimensions of posture during social interaction:

- **Non-inclusiveness - Inclusiveness**: Individuals in an interaction tend to define and limit group space by the placement of their bodies or extremities (arms and legs). If the individuals are in a line, *bookending* can be observed. The individuals at both ends turn inwards and extend their extremities across the open space.
- **Vis-à-vis - Parallel**: In an interaction, individuals can either situate themselves face to face (*vis-à-vis*) or side by side (*parallelism*) [SS72]. Individuals situate themselves vis-à-vis usually in an interaction involving an exchange of information, for instance, one-to-one quarrelling. In contrast, parallelism is used when individuals are involved in an interaction towards some third party: two individuals quarrelling against a third.
- **Non-congruence - Congruence**: The ways between which the bodies of individuals in an interaction are arranged in compliment to each other is termed *postural congruence*, and is closely related so *interactional synchrony* [Ken70] (see 6.1. Postural congruence indicates similarity in views and gives an indication of status [Sch64]. Congruent body postures can occur in interactions that are vis-à-vis or in parallel.

Posture is also important for expressing attitude [Arg98] (see section 6.2). High affiliation (liking) is expressed through postures that bring people closer (e.g. leaning forward while sitting) and by open armed postures, while low affiliation is expressed by closed posture that present a barrier between people (e.g. crossing arms). Bécheiraz and Thalmann [BT96]'s characters were able to choose different postures that displayed different levels of affiliation. Another dimension of attitude is status: dominance and submission. Dominant people tend to have large open postures, that increase their size, such as standing straight, expanding the chest and putting hands on hips, while submissive people tend to have small, closed and hunched over postures. Gillies and Ballin [GB03] used a model of both affiliation and status to model interaction between characters.

Animating posture is relatively simple. It is mostly done using a library of pre-defined postures from which an appropriate posture is chosen based on some of the factors listed above (for example the system by Guye-Vuillème *et al.* [GVCP*99]. Transitioning between postures can use a standard motion transitioning method [RGBC96]. In order to make sure there is enough variety in posture some randomness is used when choosing postures. For even more variety Gillies and Ballin [GB04] choose multiple postures at a time from the library and interpolate them to generate new postures, using random weights (see figure 2(g) and the next section for a discussion of motion interpolation). Though most methods use libraries of posture there are also procedural methods of posture generation that generate posture without the need for a library of data. For example, Densley and Willis [DW97] use a number of "posture functions" which are coordinate rotations of joints of the body. Functions can either apply to specific body parts (specific groups of joints), or to the whole body. Their model relates these posture functions to a set of basic emotions.

7.3.2. Quality of Body movement

Many researchers believe that the bodily expression of movement is a highly accurate portrayal of the emotional state of an individual. Studies using dynamic point light displays on a human body have shown that individuals are capable of accurately recognising a meaningful movement in about a quarter of a second [Joh73].

Studies were conducted by Wallbott [Wal98] and Montepare *et al.* [MKZA99] to explore the use of body movements, posture and gestures as cues to the emotional state of individuals. Wallbott [Wal98] concluded that there was a distinct pattern postural behaviour associated with at least some emotions both in the qualitative and quantitative sense. Another important factor is the level of movement activity associated with the emotions. For instance, energised movements were typical of joy, anger and fear in that order while less movement was associated with despair and shame. While Wallbott's studies [Wal98] were focused on exploring behavioural cues by asking actors to act emotional expressions out and then coding out the behaviours; Dittman *et al.* [DPB65] and Montepare *et al.* [MKZA99] investigated the extent to which individuals perceive bodily expressions of emotions. Montepare *et al.* [MKZA99], showed participants a set of three second muted video dramatisations of two actors in behavioural scenes depicting one of four emotional states: anger, sad, happy and neutral. The faces and dialogues of the actors in the clips were blurred and eliminated before being used in the study. Participants recorded the dominant emotion perceived in the clips and rating the clips with respect to characteristics of body movement on a set of six 7-point response scales: smooth-jerky, stiff-loose, soft-hard, slow-fast, expanded-contracted, and no action-lot of action [MKZA99]. Results indicated that neutral clips were identified with a higher degree of accuracy than emotional clips. Amongst the emotional clips, angry clips were identified more accurately than sad or happy clips both of which were identified with similar levels of accuracy. Angry clips were characterised by individuals to be jerky, stiff, fast, hard, expanded and full of actions. In addition, angry displays of emotion were recognised with the most accuracy [MKZA99].

Elsewhere Paterson *et al.* reported studies in which individuals were shown arm movements such as eating, drinking, lifting and knocking movements, obtained from actors posing the movements in one of 10 *emotions* including anger, sad and neutral [PPS01]. Analysis of the movements suggested a correlation between the emotion category and the velocity of the motion. Unsurprisingly sad movements were always slower than neutral while angry movements were faster than neutral. Energetic movements were positively correlated with shorter durations and greater magnitudes of average velocity, peak velocity, acceleration, deceleration, and jerk. A further data collection process confirmed that angry movements have the shortest duration and highest velocities while sad movements have the longest duration and lowest velocities [PPS01]. Paterson *et al.* carried out a categorisation study with participants who viewed the movements in form of point light stimuli. Participants were able to correctly recognise angry and sad movements even with limited cues. When the original movements were modified temporally, angry movements of various speeds were categorised as having differing intensities of anger while sad movements were categorised as angry movements when sped up. This suggests that individuals are especially heightened to recognise angry movements even if they are modified temporally. This result is in agreement with the earlier studies conducted by Montepare *et al.* [MKZA99] in which participants recognised anger from posed body movement with the highest accuracy.

7.3.3. Modelling expressive body movement

Animating expressive body movement involves taking a particular type of motion such as walking, sitting down or drinking, and applying a style to it. A style can represent many things it can be an emotional style of movement, or an individuals' personal style of movement, which might in some way reflect their personality. Thus the same animation techniques can be applied to many of the different factors discussed.

Bruderlin and Williams [BW95] do a frequency analysis on pieces of motion and are able to alter the relative importance of various frequency bands. However, these transforms do not capture the full subtlety of emotional movement. Chi *et al.* [CCZB00] have developed motion transformations based on Laban Movement Analysis (LMA) which is a method of studying, observing, describing, notating and interpreting "human movement" [Dav01]. They have implement two components of the system, Shape (the changing form the body makes in space) and Effort (how the body concentrates its exertion while performing an action). The shape component alters the amplitude of the motion in the horizontal, vertical and sagittal (side to side) planes relative to the body. The Effort transform is more complex but involves altering the continuity and timing parameters of the motion paths and adding "flourishes" such as wrist bends.

Another approach is to learn a style from one or more examples of a style and then applying it to a new, neutral motion. This was done successfully by Patterson *et al.* in point light stimuli sets of body movement [PPS01]. The examples are typically captured motion of an actor performing an action in a given style. A variety of different methods have been used to do this learning. One of the first attempts at this was by Amaya *et al.* [ABC96]. They learned two parameters that represented a style: a temporal transform and a spatial amplitude transform. More recently, researchers have tried a number of different representations of style; Brand and Hertzmann [BH00] used Hidden Markov Models and Hsu *et al.* [HPP05] use linear matrix methods. Perhaps the most sophisticated model is by Liu *et al.* [LHP05] who use a physically-based simulation of human movement which has a number of parameters that represent style. They use an optimisation method to learn the parameters from input data. The major disadvantage of all these methods is that it is not clear that the style of motion can be completely separated from the action being performed. It is not clear that what makes an angry walk angry is the same thing that make an angry gesture angry.

Rose *et al.* [RCB98] do not attempt to extract a style that is independent of the type of action but use a set of example data to build a space of different styles for a single action. They define a *verb* which consists of a set of example motions representing a single type of action in different styles. They then define a number of *adverbs* which are numerical descriptors of a style of motion and they assign adverb values to each example motion, so example x might be 90% "happy" and 20% "surprised". Given a new set of adverb values they can generate a new motion by interpolating the example motions using radial basis functions. Mukai and Kuriyama [MK05] suggest using geo-statistical interpolation to combine motions as they are better able to represent the statistical properties of motion. This type of interpolation method requires some way of capturing large amount of motion data with associated adverb values, such as Garcia-

Rojas *et al.*'s simultaneous motion capture and annotation system [GRGTV05].

7.4. Gestures

As discussed in section 6.1 gestures are intimately linked to speech. In fact some authors consider them both to arise from a single underlying representation and process [Cas00]. As such this section will discuss gesture primarily in its relationship with verbal language.

7.4.1. Types of gestures

The first obstacle to generating gesture is the huge variety of gestures we produce in day to day conversation. Many are produced on the spur of the moment and instantly forgotten. Gestures also serve many different function in speech, sometimes replacing words, sometimes complementing their meaning, and sometimes seemingly having no meaning at all. In order to generate gestures we must first break down this variety by classifying gestures into a number of fundamental types. Cassell [Cas00] classifies gestures into three basic types:

- *emblems* are standardised gestures with a clearly defined, though culturally specific meaning, for example, the "thumbs up" gesture.
- *propositional gestures* also have a clearly defined meaning though their form is not standardised. Examples include using hands to show size when saying "it was this big" or pointing to an object then a location when saying "move that there".
- *spontaneous gestures* are unplanned and produced unconsciously, in fact we normally do not remember producing them. They commonly accompany speech (they are *co-verbal*).

Cassell largely concentrates on spontaneous gestures as they form the large majority of gestures used. They are also a challenging form of gesture to study as they are unconscious and we therefore do not have conscious understanding of they use, and their meaning is less defined. Cassell [Cas00] identifies four types of spontaneous gesture:

- *Iconic* gestures represent, by their shape, some concrete feature of an object or action being discussed. An example might be bringing the hands close together to represent a thin gap.
- *Metaphoric* gestures are similar to iconic gestures in that their shape represents something that is being discussed. However, rather than directly representing a concrete, physical feature they use a metaphor to describe a feature that may be abstract. An example, might be moving the hand forward to represent the progress of time.
- *Deitic* gestures are pointing gestures that specify a location in space. They can be the standard pointing gestures with the index figure, but can also use the whole hand or other parts of the body, such as the head, to point. Deitics can be concrete, pointing to an actual object, or metaphorical, giving a conceptual location in space to an abstract idea (often used when comparing ideas).
- *Beat* gestures are small rhythmic gestures whose shape is independent of what is being discussed. They can serve to add emphasis or mark the rhythm of speech.

Ruttkay *et al.* [RPPN02] point out that individuals can also vary in the way they use gesture: the repertiore of gestures; the degree to which they use redundant gestures; the tendency to use or not use a gesture in a given situation and the characteristics of the motion. Modelling the characteristics of motion has been discussed more in section 7.3.2, in particular Badler *et al.* [BCZC00] have applied their EMOTE model to generating gesture.

7.4.2. Timing and synchrony

A gesture can be divided into a number of phases [Cas00]. Generally iconic and metaphoric gestures have three phases (triphasic), *preparation* where the hands are moved to the correct position, *stroke* where the shape is made, and *retraction* where the hands are returned to their original position. Beat and deitic gestures have two phases (biphasic), the hand is moved into the position of the gestures and out of it again. When gestures are continuously produced during speech, they tend to blend into each other, with the retraction phase of the first combining with the preparation of the second, in what is called *co-articulation* [CPB*94, KW04].

Gestures are synchronised with the words in speech to which they correspond [CPB*94]. The timing of gestures are also closely linked to the intonation of speech with the stroke occuring at the same time as stressed words, i.e. the most emphasised words. Therefore, in order to generate correctly timed gestures it is important to have a model of which words in a sentence should be emphasised. Cassell *et al.* [CPB*94] use the theme/rheme distinction (see section 6.1.2) to determine the emphasis, and therefore the occurrence of gestures, in speech. They synchronise gestures with verb phrases in the rheme section of an utterance (the section containing new information). They choose between different types of gestures depending on the words used, words that refer to a literal shape get iconics, otherwise metaphorics are generated if there is a suitable metaphor, or deitics if there is a corresponding spatial location. Otherwise beat gestures are used.

7.4.3. Gesture planning

Generating gestures consists of two stages, firstly we must determine which type of gesture to make at a given time, and then the gestures must be animated, determining the exact details of their motion (as describe in the next section). As gestures are closely associated with language gestures planning is closely tied to the way language is handled, either by parsing language to obtain information for gesture generation [CVB01] or by simultaneously generating speech and gestures. The first is discussed in detail in section 6.1.4.

For simultaneous generation, Cassell and Stone [CS99] use a single generative grammar that encompasses both speech and gesture. Their internal representation consists of a database of *lexical descriptors* which are templates for elements of a multimodal utterance. Each descriptors contain 3 types of information. The first is *syntactic* information, consisting of a phrase structured grammar, that contains both spoken words and gestures. The second is *semantic* information, a logical representation of the combined meaning of the elements. This semantic representation can be compared to information about the characters internal knowledge and the shared knowledge that has already been already revealed in the conversation and therefore to calculate the theme (already know) and rheme (new element) of the utterance. The final piece of information is the *pragmatic* information which is used to determine whether an element is appropriate in given conversational context. The fact that each lexical descriptor contains both words and gestures means that it is possible to use it in the same way as a traditional dialogue generation system but generate simultaneous gesture and speech. The original system used information about complete gestures from a *gestionary* (see next section) but this work has been extended by Kopp, Tepper and Cassell [KTC05] so that the grammar contains features of gestures from which new gestures can be generated on the fly using Kopp and Wachsmuth's system [KW00,KW04] (see next section).

7.4.4. Gesture animation

The final stage of gesture generation is animation, calculating the movements of the characters. A standard approach is to use a database of motion captured or hand animation movements called a *gestionary*. Roughly each movement is a single gesture and the animation phase consists in choosing an appropriate motion and playing it on a character. The problem with this approach is its lack of flexibility, the range of possible gestures is limited to what is in the database. This is not so much of a problem for beat gestures, but can be very problematic for other types of gestures, such as deitics as the meaning of the gesture depends on the exact direction of the gesture, so it can be confused if there is not an exactly matching gesture in the gestionary.

This problem has prompted Kopp and Wachsmuth [KW00,KW04] to use procedural animation. In their system a gesture is represented at a high level as a set of motion constraints, which are used at run time to generate the final motion. Motion generation is performed by an algorithm based on empirical data on human movement. They use a database of template gestures. A template gestures is an incomplete set of motion constraints which takes a number of parameters. These parameters are used to calculate the remaining constraints needed to generate the motion. When a gesture is needed a template is chosen and parameters supplied. The gesture is then scheduled in time and generated using based on the constraints. The fact that gestures are represented as constraints and motions is only generated at the last minute gives a lot of flexibility. In particular the generation of a gesture can take into account the motions before and after it in order to achieve co-articulation. The motion generator can also take into account timing constraints to ensure that gestures are synchronised with speech.

7.5. Proxemics

Proxemics was a term coined by Edward T Hall to encompass the interrelated observations and theories of mans' use of space in social situations [Hal69]. The flow and changes in the interpersonal distances between individuals in a conversation is an integral part of communication. If an individual get too close, the response is instantaneous and automatic, the conversational partner backs up and if an individual wants to have an intimate conversation, they drop their voice to a soft utterance thereby instigating the other to move closer [Hal59]. Hall's model consists of four concentric zones; a different level of social intimacy is required for others to enter each zone. This classification has been supported by a study by Burgess [Bur83]. The zones are:

- **Public distance:** more than 12ft, the distance between stranger that are not interacting.
- **Social distance:** 4-12ft, normal social interaction,
- **Personal distance:** 1.5-4ft, close relationships
- **Intimate distance:** less than 1.5ft, touching, intimate interaction with a spouse or children

Hall's distances are averages for middle class American adults, however, there is great variation in the personal distances people find acceptable. Culture has a strong effect, for instance, English males keep further apart that French or South American males [Som69, AT80, SS72]. Personality is also important with dominant and introverted individuals having keeping greater distances to other people. Chittaro and Serra [CS04] model proximal distances as appropriate to the extraversion score of virtual characters. Hall's model relates distance to relationship, and it seems that people with high affiliation (liking) stand or sit closer to each other [Arg69]. Leipold showed that people sat closer when praised. However, an undesired intrusion into personal space can be highly aggressive and angry people also move closer [Duc98]. Proxemics is also closely related to Argyle's *Equilibrium Threory* (see section 6.2), if people are force to stand closer than they would normally (e.g. because of crowded conditions) they compensate in other ways, for example, but reducing eye contact [AD65].

Proxemics also seems to hold between real and virtual people. People maintain a realistic distance to a virtual character in immersive virtual environments [BBB01, BBBL03] and between their avatars in multi-user virtual worlds [Jef98]. In fact Bailenson, Blascovich and Beall argue for the utilisation of behaviours especially proximal behaviour as a realism gauge for immersive virtual environment [BBBL03]. Not only is proximal behaviour an integral part social norms and communication amongst individuals, it is easily recordable in virtual environments making it a very viable and potential tool to measure user responses to stimuli in virtual environments. Despite this there has been little in the way of work on virtual characters that display proxemic behaviour, Gillies and Slater's character maintains a realistic social distance [GS05], but this work is not integrated with their earlier work on attitude [GB03]. Bickmore and Picard [BP06], have a clever implementation of proxemics in a desktop based system, they use close ups of a character to indicate close proximity when discussing intimate subjects.

8. Conclusion

We end this paper by summarising some important points to bear in mind for everyone who is creating a virtual character:

- Virtual characters are an important and powerful part of the content in virtual worlds, especially with respect to invoking user responses. However, users' response to the presentation of virtual characters may vary. Different people may have different levels of acceptance and trust towards virtual characters - section 2.
- Nonverbal communication is a vital part of creating a believable character. Nonverbal communication is the "output" of a character simulation system - section 3.
- Nonverbal behaviour has many different functions, such as expressing emotions (section 4), displaying personality (section 5), regulation conversation (6.1) and forming relationships (section 6.2).
- An emotion (section 4) is a short-term phenomena that adds meaning to the verbal and nonverbal content. In contrast a personality (section 5) is a disposition towards emotions and moods.
- A virtual character must be socially intelligent. Its nonverbal communication (and thus its emotion and to some extent personality) need to respect rules of conversation (6.1) and relationship forming (section 6.2).
- Non-verbal communication is expressed through many different modalities such as facial expression (section 7.1), gaze (section 7.2), kinesics (section 7.3) and gesture (section 7.4).

The final point to make about the diversity of nonverbal communication is that it is vital to tailor it to a given application. When choosing which models of nonverbal communication to implement many different questions must be considered. For instance, will the situations be emotionally charged or not? This will determine the emotion model needed (section 4.5). Will the character be interacting with other characters, or real people or no one? This will determine the social skills needed. Will they involve persuasion and negotiation? If so emotion (section 4.5), trust and relationship models (section 6) are important. Will the character have audible speech or just text? The difference is important from the point of view of synchronising behaviour. Will people interact with a character over a long time period? If so a consistent personality (section 5.1) and the ability to form relationships (section 6) is important otherwise a basic emotion model (section 4.5.1) and politeness (section 6.2.2) may be enough. How will the character be displayed? On a small display the face is the most important part to show so gaze (section 7.2) and facial expression (section 7.1) should be modelled, if a character has a low polygon count on the other hand facial expressions might be difficult to display so body movements (section 7.3.2) and postures (section 7.3.1) are better modalities of display.

In summary we have covered a broad range of the research and development that has gone in to creating expressive characters. Many of the models and techniques that have been described draw heavily from research in psychology and social science. We have highlighted the aspects of this research that have been adopted, but also indicate where there may be more material to incorporate. It is perhaps fair to say that in the last couple of years, it is the virtual characters community that is starting to drive the development and evaluation of models of psychological state. However, this is a diverse area of research and this diversity of research is itself a challenge to researchers in this field: each character system has been designed to investigate a particular aspect of non-verbal communication. It is not yet clear if all of this research can be integrated into a single platform with a wide variety of capabilities. Thórisson is currently taking steps to tackle part of this problem [TLPD05]

9. Acknowledgements

This work was supported by the UK projects, EQUATOR Interdisciplinary Re-search Collaboration (EPSRC Grant GR/N15986/01) and Empathic Avatars (EPSRC EP/D505542/1). We would like to thank Prof. Philip Willis and Dr. Joanna J. Bryson for their roles in facilitating the collaboration between University College London and the University of Bath.

10. Figure Credits

- Figure 1(a): Virtual Marilyn by MIRAlab, University of Geneva.
- Figure 1(b): Kaya a photorealistic female virtual human created by Alceu Baptistão, Vetor Zero, Brazil.
- Figure 1(c): Virtual Characters from the game Half-Life 2 by the Valve Corporation. Image copyright Valve Corporation. Used with permission.
- Figure 1(d): Dawn a pixie by nVIDIA Corporation.
- Figure 2(a): Woggles by the School of Computer Science, Carnegie Mellon University.
- Figure 2(b): Greta by Pelachaud *et al.* [RPP*03].
- Figure 2(c): Gandalf by Thórisson [Thó99].
- Figure 2(d): Rea by Cassell *et al.* [CBB*99].
- Figure 2(e): Spark by Vilhjálmsson [Vil05]; Copyright 2003 by MIT Medialab.

- Figure 2(f): Characters used in the studies on Paranoia by Freeman *et al.* [FSB*03] and Garau *et al.* [GSPR05] at the Department of Computer Science, University College London.
- Figure 2(g): Characters used to simulate interpersonal attitudes by Gillies *et al.* [GB04] at the Department of Computer Science, University College London.
- Figure 2(h): A participant interacting with DIANA by Lok *et al.* [JDR*05].
- Figure 2(i): Character displaying joy using the models of emotion and personality described in Egges *et al.* [EKMT04] by MIRAlab, University of Geneva.
- Figure 2(j): One of the characters used in studies [GSV*03, VGSS04] to investigate visual and behavioural realism at the Department of Computer Science, University College London.
- Figure 2(k): A character displaying an expression (smile) in two different emotional states (sad and happy) using the DER system described by Tanguy *et al.* [?] at the Department of Computer Science, University of Bath.
- Figure 2(l): Bodychat2 by Vilhjálmsson [VC98]; Copyright 1999 by MIT Medialab.

References

[ABC96] AMAYA K., BRUDERLIN A., CALVERT T.: Emotion from motion. In *Graphics Interface* (June 1996), Canadian Information Processing Society / Canadian Human-Computer Communications Society, pp. 222–229.

[AC76] ARGYLE M., COOK M.: *Gaze and mutual gaze*. Cambridge University Press, 1976.

[AD65] ARGYLE M., DEAN J.: Eye contact, distance and affiliation. *Sociometry 28* (1965), 289–304.

[AI72] ARGYLE M., INGHAM R.: Mutual gaze and proximity. *Semiotica 6* (1972), 32–49.

[AIAM73] ARGYLE M., INGHAM R., ALKEMA F., MCCALLIN M.: The different functions of gaze. *Semiotica 7* (1973), 10–32.

[AKG*99] ANDRÉ E., KLESEN M., GEBHARD P., ALLEN S., RIST T.: Integrating models of personality and emotions into lifelike characters. In *Proceedings of the Workshop on Affect in Interactions Ű Towards a new Generation of Interfaces* (1999), pp. 136–149.

[ALC74] ARGYLE M., LEFEBVRE L., COOK M.: The meaning of five patterns of gaze. *European Journal of Social Psychology 4*, 2 (1974), 125–136.

[Arg69] ARGYLE M.: *Social interaction*. Tavistock publications, 1969.

[Arg98] ARGYLE M.: *Bodily Communication*, second ed. Methuen & Co Ltd, 1998.

[Art05] ARTS E.: The sims online, http://thesims.ea.com/, January 2005.

[ASHS05] ALBRECHT I., SCHRÖDER M., HABER J., SEIDEL H.-P.: Mixed feelings: Expression of non-basic emotions in a muscle-based talking head. *Special issue of Journal of Virtual Reality on Language, Speech & Gesture 8* (2005), 201–212.

[AT80] ARGYLE M., TROWER P.: *Person to person: ways of communicating*. Harper and Row, 1980.

[Bar02] BARTNECK C.: Integrating the occ model of emotions in embodied characters. In *Proceedings of the Workshop on Virtual Conversational Characters* (2002), pp. 39–48.

[BAZB02] BADLER N., ALLBECK J., ZHAO L., BYUN M.: Representing and parameterizing agent behaviors. In *Proceeding Computer Animation* (Geneva, Switzerland, 2002), IEEE Computer Society June 2002, pp. 133–143.

[BB98] BREESE J. S., BALL G.: *Modeling Emotional State and Personality for Conversational Agents*. Tech. Rep. MSR-TR-98-41, March 1998.

[BB04] BAILENSON J. N., BLASCOVICH J.: *Avatars*, first ed. Encyclopedia of Human-Computer Interaction. Berkshire Publishing Group, September 2004, pp. 64–68.

[BBB01] BAILENSON J. N., BLASCOVICH J., BEALL A. C.: Equilibrium theory revisited: Mutual gaze and personal space in virtual environments. *Presence: Teleoperators and Virtual Environments 10*, 6 (2001), 583–598.

[BBBL03] BAILENSON J. N., BLASCOVICH J., BEALL A. C., LOOMIS J.: Interpersonal distance in immersive virtual environments. *Personality and Social Psychology Bulletin 29*, 7 (2003), 1–15.

[BC97] BAVELAS J. B., CHOVIL N.: Faces in dialogue. In *The Psychology of Facial Expression*, Russel J. A., Fernández-Dols J. M., (Eds.). Cambridge University Press, 1997, pp. 334–346.

[BC00] BAVELAS J. B., CHOVIL N.: Visible acts of meaning: An integrated message model of language in face-to-face dialogue. *Journal of Language and Social Psychology 19*, 2 (2000), 163–194.

[BC01] BICKMORE T. W., CASSELL J.: Relational agents: A model and implementation of building user trust. In *Proceedings of SIGCHI* (2001), pp. 396–403.

[BC05] BICKMORE T., CASSELL J.: Social dialogue with embodied conversational agents. In *Advances in Natural, Multimodal Dialogue Systems*, van Kuppevelt J., Dybkjaer L.and Bernsen N., (Eds.). Kluwer Academic, 2005, pp. 256–284.

[BCZC00] BADLER N., COSTA M., ZHAO L., CHI D.: To gesture or not to gesture: What is the question? In *Computer Graphics International (CGI)* (June 2000), pp. 3–9.

[BH00] BRAND M., HERTZMANN A.: Style machines. 183–192.

[BH04] BREWER M. B., HEWSTONE M.: *Emotion and motivation*. Blackwell Publishing, 2004.

[Bir68] BIRDWHISTELL R. L.: Kinesics. *International encyclopaedia of social sciences 8* (1968), 379–385.

[Bir71] BIRDWHISTELL R. L.: *Kinesics and context: essays on body motion communication*. Allen Lane The Penguin Press, 1971.

[BKW04] BECKER C., KOPP S., WACHSMUTH I.: Simulating the emotion dynamics of a multimodal conversational agent. In *Affective Dialogue Systems* (June 2004), pp. 154–165.

[BL78] BROWN P., LEVINSON S.: Universals in language usage: Politeness phenomena. In *Questions and politeness*, Goody E., (Ed.). Cambridge University Press, 1978.

[BLB*02] BLASCOVICH J., LOOMIS J., BEALL A. C., SWINTH K., HOYT C., BAILENSON J. N.: Immersive virtual environment technology as a methodological tool for social psychology. *Psychology Inquiry 13* (2002), 103–124.

[BLR94] BATES J., LOYALL B. A., REILLY S. N.: An architecture for action, emotion and social behaviour. In *Selected papers from the 4th European Workshop on on Modelling Autonomous Agents in a Multi-Agent World, Artificial Social Systems* (1994), Springer-Verlag, pp. 55–68.

[BP04] BICKMORE T. W., PICARD R. W.: Towards caring machines. In *ACM SIGCHI extended abstracts on Human factors in computing systems* (2004), pp. 1489–1492.

[BP06] BICKMORE T., PICARD R.: Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer Human Interaction (ToCHI)* (2006).

[BR05] BARTNECK C., REICHENBACH J.: Subtle emotional expressions of synthetic characters. *International journal of human-computer studies 62*, 2 (February 2005), 179–192.

[BT96] BÉCHEIRAZ P., THALMANN D.: A model of nonverbal communication and interpersonal relationship between virtual actors. *Computer Animation* (June 1996), 58–67.

[Bui04] BUI T. D.: *Creating Emotions And Facial Expressions For Embodied Agents*. PhD thesis, 2004.

[Bul87] BULL P. E.: *Posture and Gesture*, vol. 16 of *International Series in Experimental Social Psychology*. Pergamon Press, 1987.

[Bur83] BURGESS W. J.: Interpersonal spacing behavior between surrounding nearest neighbors reflects both familiarity and environmental density. *Ethology and Sociobiology 4*, 1 (1983), 11–17.

[BW95] BRUDERLIN A., WILLIAMS L.: Motion signal processing. *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques (SIGGRAPH)* (August 1995), 97–104.

[Can15] CANNON W. B.: *Bodily changes in pain, hunger, fear and rage*. Appleton, 1915.

[Cas00] CASSELL J.: Nudge nudge wink wink: Elements of face-to-face conversation for embodied conversational agents. In *Embodied Conversational Agents*, Cassell J., Sullivan J., Prevost S.,, Churchill E., (Eds.). MIT Press, 2000, pp. 1–27.

[CB01] CASANUEVA J. S., BLAKE E. H.: *The effects of avatars on co-presence in a collaborative virtual environment*. Tech. Rep. CS01-02-00, 2001.

[CB03] CASSELL J., BICKMORE T.: Negotiated collusion: Modeling social language and its relationship effects in intelligent agents. *User Modeling and User-Adapted Interaction 13*, 1-2 (February 2003), 89–132.

[CBB*99] CASSELL J., BICKMORE T. W., BILLINGHURST M., CAMPBELL L., CHANG K., VILHJÁLMSSON H. H., YAN H.: Embodiment in conversational interfaces: Rea. In *Proceedings of SIGCHI* (1999), ACM Press, pp. 520–527.

[CBC*99] CASSELL J., BICKMORE T. W., CAMPBELL L., CHANG K., VILHJÁLMSSON H. H., YAN H.: Requirements for an architecture for embodied conversational characters. In *Computer Animation and Simulation* (1999), pp. 109–120.

[CBVY00] CASSELL J., BICKMORE T. W., VILHJÁLMSSON H. H., YAN H.: More than just a pretty face: Affordances of embodiment. In *Proceedings of the 5th international conference on Intelligent user interfaces (IUI)* (2000), pp. 52–59.

[CCD00] COLBURN A. R., COHEN M. F., DRUCKER S. M.: *The Role of Eye Gaze in Avatar Mediated Conversational Interfaces*. Tech. Rep. MSR-TR-2000-81, July 2000.

[CCM02] CAVAZZA M., CHARLES F., MEAD S. J.: Character-based interactive storytelling. In *IEEE Intelligent systems* (2002), pp. 17–24.

[CCZB00] CHI D., COSTA M., ZHAO L., BADLER N. I.: The emote model for effort and shape. In *Proceedings of SIGGRAPH* (July 2000), vol. ACM Computer Graphics Annual Conference, pp. 173–182.

[Cho99] CHOPRA S.: *Where To Look?: Automating Some Visual Attending Behaviours of Human Characters*. PhD thesis, June 1999.

[CNB*01] CASSELL J., NAKANO Y. I., BICKMORE T. W., SIDNER C. L., RICH C.: Non-verbal cues for discourse structure. In *Proceedings of the 39th Annual Meeting on Association for Computational Linguistics* (2001), Association for Computational Linguistics, pp. 114–123.

[Col85] COLLIER G.: *Emotional Expression*. Lawrence Erlbaum Associates, 1985.

[Cou04] COULSON M.: Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. *Journal of Nonverbal Behavior 28*, 2 (2004), 117–139.

[CPB*94] CASSELL J., PELACHAUD C., BADLER N. I., STEEDMAN M., ACHORN B., BECKET T., DOUVILLE B., PREVOST S., STONE M.: Animated conversation: Rule-based generation of facial expression, gesture & spoken intonation for multiple conversational agents. In *Proceedings of SIGGRAPH* (1994).

[CR96] CARROLL J. M., RUSSELL J. A.: Do facial expressions signal specific emotions? judging emotion from the face in context. *Journal of personality and social psychology 70*, 2 (1996), 205–218.

[CS99] CASSELL J., STONE M.: Living hand to mouth: Psychological theories about speech and gesture in interactive dialogue systems. In *Proceedings of the AAAI Fall Symposium on Psychological Models of Communication in Collaborative Systems* (November 1999), pp. 34–42.

[CS04] CHITTARO L., SERRA M.: Behavioral programming of autonomous characters based on probabilistic automata and personality. *Journal of Visualization and Computer Animation 15*, 3-4 (2004), 319–326.

[CS05] COWELL A. J., STANNEY K. M.: Manipulation of non-verbal interaction style and demographic embodiment to increase anthropomorphic computer character credibility. *International journal of human-computer studies 62*, 2 (February 2005), 281–306.

[CV99] CASSELL J., VILHJÁLMSSON H. H.: Fully embodied conversational avatars: Making communicative behaviours autonomous. In *International Conference on Autonomous Agents and multiagent systems* (1999), vol. 2, pp. 45–64.

[CVB01] CASSELL J., VILHJÁLMSSON H. H., BICKMORE T. W.: Beat: The behaviour expression animation toolkit. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques (SIGGRAPH)* (2001), pp. 477–486.

[Dam95] DAMASIO A. R.: *Descartes Error : Emotion, Reason, and the Human Brain*. Pan Macmillan, April 1995.

[Dar72] DARWIN C.: *The Expression of the Emotions in Man and Animals (Orignially published in 1872, reprinted in 1999)*, third ed. Fontana Press, April 1872.

[Dav01] DAVIES E.: *Beyond Dance, Laban's Legacy of Movement Analysis*. Berchin Books, 2001.

[DCCP02] DE CAROLIS B., CAROFIGLIO V., PELACHAUD C.: From discourse plans to believable behavior generation. In *Proceedings of the International Natural Language Generation Conference* (July 2002), pp. 65–72.

[DLN03] DENG Z., LEWIS J. P., NEUMANN U.: Practical eye movement model using texture synthesis. In *In Proceedings of ACM SIGGRAPH Conference on Sketches and Applications* (July 2003).

[DLN05] DENG Z., LEWIS J. P., NEUMANN U.: Automated eye motion using texture synthesis. *IEEE Computer Graphics and Applications 25*, 2 (May 2005), 24–30.

[DLS*89] DUCLOS S. E., LAIRD J. D., SCHNEIDER E., SEXTER M., STERN L., VAN LIGHTEN O.: Emotion-specific effects of facial expressions and postures on emotional experience. *Journal of Personality and Social Psychology 57*, 1 (July 1989), 100–108.

[Doh97] DOHERTY W.: The emotional contagion scale: A measure of individual differences. *Journal of Nonverbal Behaviour 21*, 2 (1997), 131–154.

[DPB65] DITTMANN A. T., PARLOFF M. B., BOOMER D. S.: Facial and bodily expression: A study of receptivity of emotional cues. *Psychiatry 28* (1965), 239–244.

[DRPP04] DE ROSIS F., POGGI I., PELACHAUD C.: Transcultural believability in embodied agents: a matter of consistent adaptation. In *Agent Culture: Designing Virtual Characters for a Multi-Cultural World*, Trappl R., Petta P., (Eds.). Lawrence Erlbaum Associates, 2004.

[DRS02] DECARLOS D., REVILLA C., STONE M.: Making discourse visible: Coding and animating conversational facial displays. In *Computer Animation 2002* (2002), IEEE Computer Society, pp. 11–16.

[DSKBB05] DE SILVA R. P., KELSMITH A., BIANCHI-BERTHOUZE N.: Towards unsupervised detection of affective state nuances from body postures. In *Proceedings of the First International on Conference on Affective Computing & Intelligent Interaction* (October 2005), Tao J., Tan T.,, Picard R. W., (Eds.), vol. 3784 of *Lecture Notes in Computer Science*, Springer, pp. 32–39.

[Duc98] DUCK S.: *Human relationships*, third ed. SAGE, 1998.

[DW97] DENSLEY D. J., WILLIS P. J.: Emotional posturing: A method towards achieving emotional figure animation. In *Computer Animation* (June 1997), IEEE, Piscataway, NJ, USA, pp. 8–14. Compilation and indexing terms, Copyright 2004 Elsevier Engineering Information, Inc.

[ED94] EKMAN P., DAVIDSON R. J.: *The nature of emotion: fundamental questions*. Series in affective science. Oxford Unviersity Press, 1994.

[EE85] EYSENCK H., EYSENCK M.: *Personality and individual differences: a natural science approach*. Plenum Press, 1985.

[EF67] EKMAN P., FRIESEN V. W.: Head and body cues in the judgement of emotion: A reformulation. *Perceptual and Motor Skills 24* (1967), 711–724.

[EF69] EKMAN P., FRIESEN W. V.: The repertoire of nonverbal behaviour: categories, origins, usage and coding. *Semiotica 1* (1969), 49–98.

[EF75] EKMAN P., FRIESEN W. V.: *Unmasking the face: A guide to recognizing emotions from facial cues*. Prentice-Hall, 1975.

[EF76] EKMAN P., FRIESEN W. V.: *Pictures of facial affects*. Consulting psychologist press, Palo Alto, 1976.

[EF78] EKMAN P., FRIESEN W. V.: *Facial Action Coding System*. Consulting Psychologist Press, Inc., 1978.

[Ekm65] EKMAN P.: *Communication through nonverbal behaviour: A source of information about an interpersonal relationship*. Affect, cognition and personality: empirical studies eds. S. S. Tomkins and C. E. Izard. Springer publishing company, inc., 1965, ch. XIII, pp. 390–442.

[Ekm79] EKMAN P.: *About brows: Emotional and conversational signals*. Human ethology: Claims and limits of a new discipline. Cambridge University Press, 1979, pp. 169–248.

[Ekm82] EKMAN P.: *Emotion in the Human Face*, vol. Second. Cambridge University Press, New York, 1982.

[Ekm92] EKMAN P.: Facial expressions of emotion: new findings, new questions. *Psychological science 3*, 1 (1992), 34–38.

[Ekm99] EKMAN P.: *Facial Expressions*. John Wiley & Sons Ltd., 1999, ch. 16.

[EKMT02] EGGES A., KSHIRSAGAR S., MAGNENAT-THALMANN N.: Imparting individuality to virtual humans. In *First International Workshop on Virtual Reality Rehabilitation (Mental Health, Neurological, Physical, Vocational)* (November 2002).

[EKMT04] EGGES A., KSHIRSAGAR S., MAGNENAT-THALMANN N.: Generic personality and emotion simulation for conversational agents. *Computer Animation and Virtual Worlds 15*, 1 (January 2004), 1–13.

[EMT05] EGGES A., MAGNENAT-THALMANN N.: Emotional communicative body animation for multiple characters. In *Proceedings of the First International Workshop on Crowd Simulation* (November 2005), pp. 31–40.

[ENIY99a] EL-NASR M. S., IOERGER T. R., YEN J.: Peteei: a pet with evolving emotional intelligence. In *Proceedings of the third annual conference on Autonomous Agents* (1999), ACM Press, pp. 9–15.

[ENIY99b] EL-NASR M. S., IOERGER T. R., YEN J.: A web of emotions. In *Proceedings of the Workshop on Emotion based Agent Architecture (in conjunction with Autonomous Agents)* (1999).

[ENYI00] EL-NASR M. S., YEN J., IOERGER T. R.: Flame - fuzzy logic adaptive model of emotions. In *Autonomous Agents and Multi-Agent Systems* (2000), vol. 3, Kluwer Academic Publishers., pp. 219–257.

[EW65] EXLINE R. V., WINTERS L. C.: *Affective relations and mutual glances in dyads*. Affect, cognition and personality: empirical studies eds. S. S. Tomkins and C. E. Izard. Springer publishing company, inc., 1965, ch. XI, pp. 319–350.

[FB87] FORGAS J. P., BOWER G. H.: Mood effects on person perception judgements. *Journal of personality and social psychology 53* (1987), 53–60.

[FJLC99] FLACK JR W. F., LAIRD J. D., CAVALLARO L. A.: Separate and combined effects of facial expressions and bodily postures on emotional feelings. *European Journal of Social Psychology 29*, 2 (1999), 203–217.

[FOM*02] FUKAYAMA A., OHNO T., MUKAWA N., SAWAKI M., HAGITA N.: Messages embedded in gaze of interface agents — impression management with agent's gaze. In *Proceedings of SIGCHI* (2002), pp. 41–48.

[For90] FORGAS J. P.: Affective influences on individual and group judgements. *European journal of social psychology 20* (1990), 441–453.

[FR84] FEHR B., RUSSELL J. A.: Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General 113*, 3 (1984), 464–486.

[Fri88] FRIJDA N. H.: The laws of emotion. *American Psychologist 43*, 5 (May 1988), 349–358.

[FSB*03] FREEMAN D., SLATER M., BEBBINGTON P., GARETY P. A., KUIPERS E., FOWLER D., MET A., READ C. M., JORDAN J., VINAYAGAMOORTHY V.: Can virtual reality be used to investigate persecutory ideation? *Journal of Nervous and Mental Disease 191*, 8 (August 2003), 509–514.

[FSO*01] FUKAYAMA A., SAWAKI M., OHNO T., MURASE H., HAGITA N., MUKAWA N.: Expressing personality of interface agents by gaze. In *Proceedings of INTERACT* (2001), pp. 793–794.

[Fus02] FUSSELL S. R.: *The verbal communication of emotions: interdisciplinary perspectives*. Lawrence Erlbaum Associates, 2002.

[GB03] GILLIES M., BALLIN D.: A model of interpersonal attitude and posture generation. In *Proceedings of the Fourth Workshop on Intelligent Virtual Agents (Lecture Notes in Computer Science)* (September 2003), Rist T., Aylett R., Ballin D.,, Rickel J., (Eds.), vol. 2792 of *Lecture Notes in Computer Science*, Springer, pp. 88–92.

[GB04] GILLIES M., BALLIN D.: Integrating autonomous behaviour and user control for believable agents. In *Proceedings of the Third international joint* *conference on Autonomous Agents and Multi-Agent Systems (AAMAS)* (July 2004), IEEE Computer Society, pp. 336–343.

[GBA75] GRAHAM J. A., BITTI R. P., ARGYLE M.: A cross-cultural study of the communication of emotion by facial and gestural cues. *Journal of Human Movement Studies* (1975), 178–182.

[GCB04] GILLIES M., CRABTREE B., BALLIN D.: Expressive characters and a text chat interface. In *AISB workshop on Language, Speech and Gesture for Expressive Characters* (University of Leeds, March 2004), Olivier P., Aylett R., (Eds.).

[GD02] GILLIES M., DODGSON N.: Eye movements and attention for behavioural animation. *Journal of Visualization and Computer Animation (Special issue on Graphical Autonomous Virtual Humans) 13*, 5 (2002), 287–300.

[Geb05] GEBHARD P.: Alma: a layered model of affect. In *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems* (2005), ACM Press, pp. 29–36.

[GK97] GUMP B. B., KULIK J. A.: Stress, affiliation, and emotional contagion. *Journal of Personality and Social Psychology 72*, 2 (February 1997), 305–319.

[GM01] GRATCH J., MARSELLA S.: Tears and fears: modeling emotions and emotional behaviors in synthetic agents. In *Proceedings of the fifth international conference on Autonomous agents* (2001), ACM Press, pp. 278–285.

[GM04] GRATCH J., MARSELLA S.: A domain independent framework for modeling emotion. *Journal of Cognitive Systems Research 5*, 4 (2004), 269–306.

[GM05] GRATCH J., MARSELLA S.: Lessons from emotion psychology for the design of lifelike characters. *Applied Artificial Intelligence (special issue on Educational Agents - Beyond Virtual Tutors) 19*, 3-4 (2005), 215–233.

[Gof63] GOFFMAN E.: *Behavior in public places: notes on the social organization of gatherings*. Free Press, 1963.

[Gra00] GRATCH J.: Émile: Marshalling passions in training and education. In *Proceedings of the 4th International conference of Autonomous Agents* (June 2000), pp. 325–332.

[GRA*02] GRATCH J., RICKEL J., ANDRÉ E., CASSELL J., PETAJAN E., BADLER N. I.: Creating interactive virtual humans: Some assembly required. *IEEE Intelligent Systems 17*, 4 (July 2002), 54–63.

[GRGTV05] GARCIA-ROJAS A., GUTIERREZ M., THALMANN D., VEXO F.: Multimodal authoring tool for populating a database of emotional reactive animations. In *Proceedings of the 2nd Joint Workshop on Multimodal Interaction and Related Machine Learning Algorithms* (July 2005).

[GRW05] GUPTA S., ROMANO D. R., WALKER M. A.: Politeness and variation in synthetic social interaction. In *H-ACI Human-Animated Characters Interaction Workshop in conjunction with the 19th British HCI Group Annual Conference* (2005), Romano D., (Ed.).

[GS05] GILLIES M., SLATER M.: Non-verbal communication for correlational characters. In *Proceedings of the 8th Annual International Workshop on Presence* (September 2005), Slater M., (Ed.).

[GSBS01] GARAU M., SLATER M., BEE S., SASSE M. A.: The impact of eye gaze on communication using humanoid avatars. In *Proceedings of SIGCHI* (March 2001), pp. 309–316.

[GSPR05] GARAU M., SLATER M., PERTAUB D.-P., RAZZAQUE S.: The responses of people to virtual humans in an immersive virtual environment. *Presence: Teleoperators and Virtual Environments 14*, 1 (February 2005), 104–116.

[GSV*03] GARAU M., SLATER M., VINAYAGAMOORTHY V., BROGNI A., STEED A., SASSE A. M.: The impact of avatar realism and eye gaze control on the perceived quality of communication in a shared immersive virtual environment. In *Proceedings of SIGCHI* (April 2003), pp. 529–536.

[GSW*01] GUSTAVSSON C., STRINDLUND L., WIKNERTZ E., BEARD S., HUYNH Q., MARRIOTT A., STALLO J.: Vhml. Working Draft v0.3 October 21st 2001, 10 2001.

[GVCP*99] GUYE-VUILLÈME A., CAPIN K. T., PANDZIC S. I., MAGNENAT-THALMANN N., THALMANN D.: Nonverbal communication interface for collaborative virtual environments. *The Virtual Reality Journal 4* (1999), 49–59.

[GW93] GROSSMAN M., WOOD W.: Sex differences in intensity of emotional experience: a social role interpretation. *Journal of Personality and Social Psychology 65*, 5 (November 1993), 1010–1022.

[Hal59] HALL E. T.: *The silent language*, first ed. Doubleday, 1959.

[Hal69] HALL E. T.: *The hidden dimension: man's use of space in public and private*. Bodley Head, 1969.

[HBK97] HESS U., BLAIRY S., KLECK R. E.: The intensity of emotional facial expressions and decoding accuracy. *Journal of Nonverbal Behaviour 21*, 4 (1997), 241–257.

[HBR*04] HERBELIN B., BENZAKI P., RIQUIER F., RENAULT O., THALMANN D.: Using physiological measures for emotional assessment: a computer-aided tool for cognitive and behavioural therapy. In *Proceedings of ICDVRAT 2004, 5th International Conf on Disability, Virtual Reality* (2004), pp. 307–314.

[HCR94] HATFIELD E., CACIOPPO J. J., RAPSON R.: *Emotional contagion*. Cambridge Unviersity Press, 1994.

[HEG98] HARTUNG F., EISERT P., GIROD B.: Digital watermarking of (mpeg-4) facial animation parameters. *Computers and Graphics 22*, 4 (1998), 425–435.

[HMP] HARTMANN B., MANCINI M., PELACHAUD C.:.

[HPP05] HSU E., PULLI K., POPOVIĆ J.: Style translation for human motion. *Proceedings of ACM SIGGRAPH 24*, 3 (2005), 1082–1089.

[HW05] HALL L., WOODS S.: *Empathic interaction with synthetic characters: the importance of similarity*. Encyclopedia of Human-Computer Interaction. Idea Group Publishing, December 2005.

[HWA*05] HALL L., WOODS S., AYLETT R., NEWALL L., PAIVA A.: Achieving empathic engagement through affective interaction with synthetic characters. In *Proceedings of the First International Conference on Affective Computing and Intelligent Interaction* (October 2005), Tao J., Tan T.,, Picard R. W., (Eds.), vol. 3784, Springer-Verlag, pp. 731–738.

[IDN87] ISEN A. M., DAUBMAN K. A., NOWICKI G. P.: Positive affect facilitates creative problem solving. *Journal of personality and social psychology 52*, 6 (1987), 1122–1131.

[IN00] ISBISTER K., NASS C.: Consistency of personality in interactive characteristics: verbal cues, non-verbal cues and user characteristics. *International Journal of Human-Computer Studies 53* (2000), 251–267.

[Ise87] ISEN A. M.: *Positive affect, cognitive processes and social behaviour*, vol. 20 of *Advances in experimental social psychology*. Academic Press, 1987, pp. 203–253.

[Iza71] IZARD C. E.: *The face of emotion*. Appleton-Century-Crofts, 1971.

[Iza97] IZARD C. E.: Emotions and facial expression: A perspective from differential emotions theory. In *The Psychology of Facial Expression*, Russel J. A., Fernández-Dols J. M., (Eds.). Cambridge University Press, 1997, pp. 57–77.

[Jac03] JACKSON P.: The lord of the rings: The trilogy. 16mm 548 min, New Line Cinema, 2003.

[Jam32] JAMES W. T.: A study of the expression of bodily posture. *Journal of general psychology 42* (1932), 90–94.

[JDR*05] JOHNSEN K., DICKERSON R., RAIJ A., LOK B., JACKSON J., SHIN M., HERNANDEZ J., STEVENS A., LIND D. S.: Experiences in using immersive virtual characters to educate medical communication skills. In *Proceedings of the IEEE Virtual Reality Conference* (2005), vol. 324, pp. 179–186.

[Jef98] JEFFREY P.: Personal space in a virtual community (student poster). In *Proceedings of SIGCHI* (April 1998), pp. 347–348.

[Joh73] JOHANSSON G.: Visual perception of biological motion and a model for its analysis. *Perception and psychophysics 14* (1973), 201–211.

[JT05] JAN D., TRAUM D. R.: Dialog simulation for background characters. In *Proceedings of the 5th International Workshop on Intelligent Virtual Agents (Lecture Notes in Computer Science)* (September 2005), Panayiotopoulos T., Gratch J., Aylett R., Ballin D., Olivier P.,, Rist T., (Eds.), vol. 3661, Springer, pp. 65–74.

[KAPN00] KESTEREN A.-J. V., AKKER R. O. D., POEL M., NIJHOLT A.: Simulation of emotions of agents in virtual environments using neural networks. In *Learning to Behave: Internalising Knowledge. Twente Workshop on Language Technology 18* (November 2000), pp. 137–147.

[KB96] KELTNER D., BUSWELL B. N.: Evidence for the distinctness of embarrassment, shame, and guilt: A study of recalled antecedents and facial expressions of emotion. *Cognition and Emotion 10*, 2 (1996), 155–171.

[KDSBB05] KELSMITH A., DE SILVA R. P., BIANCHI-BERTHOUZE N.: Grounding affective dimensions into posture features. In *Proceedings of the First International on Conference on Affective Computing & Intelligent Interaction* (October 2005), Tao J., Tan T.,, Picard R. W., (Eds.), vol. 3784 of *Lecture Notes in Computer Science*, Springer, pp. 263–270.

[Ken67] KENDON A.: Some functions of gaze-direction in social interaction. *Acta Psychologica 26* (1967), 22–63.

[Ken70] KENDON A.: Movement coordination in social interaction: some examples described. *Acta Psychologica 32*, 2 (1970), 1–25.

[Ken90] KENDON A.: *Conducting Interaction: Patterns of Behavior in Focused Encounters*. Studies in Interactional Sociolinguistics. Cambridge University Press, 1990.

[KK81] KLEINGINNA P., KLEINGINNA A.: A categorized list of emotion definitions with suggestions for a consensual definition. *Motivation and Emotion 5* (1981), 345–379.

[Kle86] KLEINKE C. L.: Gaze and eye contact: a research review. *Presence: Teleoperators and Virtual Environments 100*, 1 (1986), 78–100.

[KMMTT91] KALRA P., MANGILI A., MAGNENAT-THALMANN N., THALMANN D.: Smile: A multilayered facial animation system. In *Proceedings IFIP conference on Modelling in Computer Graphics* (1991), L K. T., (Ed.), Springer, pp. 189–198.

[KMP02] KLEIN J., MOON Y., PICARD R. W.: This computer responds to user frustration: Theory, design and results. *Interacting with Computers 14*, 2 (2002), 119–140.

[KMT02a] KSHIRSAGAR S., MAGNENAT-THALMANN N.: A multilayer personality model. In *Proceedings of the*

2nd International Symposium on Smart Graphics (June 2002), pp. 107–115.

[KMT02b] KSHIRSAGAR S., MAGNENAT-THALMANN N.: Virtual humans personified. In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems* (July 2002), ACM Press, pp. 356–357.

[KMTGV*02] KSHIRSAGAR S., MAGNENAT-THALMANN N., GUYE-VUILLÈME A., THALMANN D., KAMYAB K., MAMDANI E.: Avatar markup language. In *Proceedings of the Eurographics workshop on Virtual environments* (2002), Eurographics Association, pp. 169–177.

[KSS96] KURLANDER D., SKELLY T., SALESIN D.: Comic chat. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques (SIGGRAPH)* (1996), ACM Press, pp. 225–236.

[KTC05] KOPP S., TEPPER P., CASSELL J.: Towards integrated microplanning of language and iconic gesture for multimodal output. In *Proceedings of the International Conference on Multimodal Interfaces (ICMI)* (2005), ACM Press, pp. 97–104.

[KVVHJ05] KIM Y., VAN VELSEN M., HILL JR. R. W.: Modeling dynamic perceptual attention in complex virtual environments. In *Proceedings of the 5th International Workshop on Intelligent Virtual Agents (Lecture Notes in Computer Science vol. 3661)* (September 2005), Panayiotopoulos T., Gratch J., Aylett R., Ballin D., Olivier P.,, Rist T., (Eds.), vol. 3661, Springer-Verlag, pp. 371–380.

[KW00] KOPP S., WACHSMUTH I.: A knowledge-based approach for lifelike gesture animation. In *Proceedings of the 14th European Conference on Artificial Intelligence (ECAI)* (2000), Horn W., (Ed.), IOS Press, pp. 663–667.

[KW04] KOPP S., WACHSMUTH I.: Synthesizing multimodal utterances for conversational agents. *The Journal Computer Animation and Virtual Worlds 15*, 1 (2004), 39–52.

[LAAB02] LATTA C., ALVARADO N., ADAMS S. S., BURBECK S.: An expressive system for endowing robots or animated characters with affective facial displays. In *The 2002 conference of the British Society for Artificial Intelligence and the Simulation of Behavior* (2002).

[LAJ05] LIM M. Y., AYLETT R., JONES C. M.: Affective guide with attitude. In *Proceedings of the First International Conference on Affective Computing and Intelligent Interaction* (October 2005), Tao J., Tan T.,, Picard R. W., (Eds.), vol. 3784, Springer-Verlag, pp. 772–779.

[Las87] LASSETER J.: Principles of traditional animation applied to 3d computer animation. *ACM SIGGRAPH Computer Graphics 21*, 4 (July 1987), 35–44.

[Laz91] LAZARUS R. S.: *Emotion and Adaptation*. Oxford University Press, 1991.

[LBB02] LEE S. P., BADLER N. I., BADLER J.: Eyes alive. In *SIGGRAPH* (July 2002), pp. 637–644.

[Lev94] LEVENSON R. W.: *Human emotion: A functional view*. The nature of emotion: fundamental questions. Oxford University Press, 1994, pp. 123–126.

[LHP05] LIU C. K., HERTZMANN A., POPOVIĆ Z.: Learning physics-based motion style with nonlinear inverse optimization. *Proceedings of ACM SIGGRAPH 24*, 3 (2005), 1071–1081.

[Lin01] LINKLATER R.: Waking life. 16mm 99 min, Fox Searchlight Pictures, 2001.

[LN02] LISETTI C. L., NASOZ F.: Maui: a multimodel affective user interface. In *Multimedia* (December 2002), pp. 161–170.

[LWB00] LANGTON S. R. H., WATT R. J., BRUCE V.: Do the eyes have it? cues to the direction of social attention. *Trends in Cognitive Sciences 4*, 2 (February 2000), 50–59.

[MB99] MARK G., BECKER B.: Designing believable interaction by applying social conventions. In *Applied Artificial Intelligence* (1999), vol. 13, pp. 297–320.

[MBP05] MANCINI M., BRESIN R., PELACHAUD C.: From acoustic cues to expressive ecas. In *The 6th International Workshop on Gesture in Human-Computer Interaction and Simulation* (2005), Lecture Notes in Artificial Intelligence, Springer.

[MC03] MIGNAULT A., CHAUDHURI A.: The many faces of a neutral face: head tilt and perception of dominance and emotion. *Journal of Nonverbal Behavior 27*, 2 (2003), 111–132.

[MD04] MULLINS D. T., DUKE M. P.: Effects of social anxiety on nonverbal accuracy and response time i: facial expressions. *Journal of nonverbal behaviour 28*, 1 (2004), 3–33.

[Meh71] MEHRABIAN A.: *Silent Messages*, first ed. Wadsworth Publishing Company, January 1971.

[Meh80] MEHRABIAN A.: *Basic dimensions for a general psychological theory: Implications for personality, social, environmental, and developmental studies*. Oelgeschlager, Gunn & Hain, 1980.

[Meh96a] MEHRABIAN A.: Analysis of the big-five personality factors in terms of the pad temperament model. *Australian Journal of Psychology 48*, 2 (1996), 86–92.

[Meh96b] MEHRABIAN A.: Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology 14* (1996), 261–292.

[MF69] MEHRABIAN A., FIRAR J. T.: Encoding of attitude by a seated communicator via posture and position cues. *Journal of Consulting and Clinical Psychology 33*, 3 (1969), 330–336.

[MG79] Muirhead R. D., Goldman M.: Mutual eye contact as affected by seating position, sex and age. *Journal of Social Psychology 109* (1979), 201–206.

[MG01] Marsella S., Gratch J.: Modeling the interplay of emotions and plans in multi-agent simulations. In *Cognitive Science Society* (2001), vol. Proceedings of the 23rd Annual Conference Cognitive Science Society.

[MG02] Marsella S., Gratch J.: A step toward irrationality: using emotion to change belief. In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems* (2002), ACM Press, pp. 334–341.

[MGM05] Maatman R. M., Gratch J., Marsella S.: Natural behavior of a listening agent. In *Proceedings of the 5th International Working Conference on Intelligent Virtual Agents* (September 2005), Panayiotopoulos T., Gratch J., Aylett R., Ballin D., Olivier P.,, Rist T., (Eds.), vol. 3661 of *Lecture Notes in Computer Science*, Springer, pp. 25–36.

[Min88] Minsky M.: *The society of mind*. Touchstone, 1988.

[MJ92] McCrae R. R., John O. P.: An introduction to the five-factor model and its applications. *Journal of Personality 60* (1992), 175–215.

[MK05] Mukai T., Kuriyama S.: Geostatistical motion interpolation. *Proceedings of ACM SIGGRAPH 24*, 3 (2005), 1062–1070.

[MKZA99] Montepare J., Koff E., Zaitchik D., Albert M.: The use of body movements and gestures as cues to emotions in younger and older adults. *Journal of Nonverbal Behaviour 23*, 2 (1999), 133–152.

[MLL00] Marsella S., Lewis J. W., LaBore C.: Interactive pedagogical drama. In *Proceedings of Autonomous Agents* (2000), vol. 4th International Conference, pp. 301–308.

[MLP04] Maya V., Lamolle M., Pelachaud C.: Influences on embodied conversational agent's expressivity. In *AISB workshop on Language, Speech and Gesture for Expressive Characters* (March 2004), Aylett R., Olivier P.,, Cavazza M., (Eds.), pp. 75–85.

[Mof97] Moffat D.: Personality parameters and programs. In *Creating Personalities for Synthetic Actors* (1997), Springer-Verlag, pp. 120–165.

[MPP01] Machado I., Paiva A., Prada R.: Is the wolf angry or... just hungry? In *Proceedings of the fifth international conference on Autonomous agents* (2001), ACM Press, pp. 370–376.

[MSWB87] Mulac A., Studley L. B., Wiemann J. M., Bradac J.: Male/female gaze in same-sex and mixed-sex dyads: gender-linked differences and mutual influence. *Human communication research 13* (1987), 323–343.

[MVS*02] Mortensen J., Vinayagamoorthy V., Slater M., Steed A., Whitton M. C., Lok B.: Collaboration in tele-immersive environments. In *Proceedings of the Eight Eurographics Workshop on Virtual Environments* (May 2002), pp. 93–101.

[NB03] Nowak K. L., Biocca F.: The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators and Virtual Environments 12*, 5 (2003), 481–494.

[NIL01] Nass C., Isbister K., Lee E.: *Truth is beauty: researching embodied conversational agents*. MIT Press, 2001.

[NRSC03] Nakano Y. I., Reinstein G., Stocky T., Cassell J.: Towards a model of face-to-face grounding. In *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics (ACL)* (2003), Association for Computational Linguistics, pp. 553–561.

[NSS01] Neumann R., Seibt B., Strack F.: The influence of mood on the intensity of emotional responses: disentangling feeling and knowing. *Cognition and Emotion 15*, 6 (2001), 725–747.

[OCC88] Ortony A., Clore G., Collins A.: *Cognitive structure of emotions*. Cambridge University Press, 1988.

[OJL87] Oatley K., Johnson-Laird P. N.: Towards a cognitive theory of emotions. *Cognition and Emotion 1* (1987), 29–50.

[Ort01] Ortony A.: *On making believable emotional agents believable*. Cambridge: MIT Press, 2001.

[Osg66] Osgood C. E.: Dimensionality of the semantic space for communication via facial expressions. *Scandanavian journal of psychology 7* (1966), 1–30.

[OT90] Ortony A., Turner T. J.: What's basic about basic emotions? *Psychological Review 97*, 3 (1990), 315–331.

[Par72] Parke Frederic I.: Computer generated animation of faces. In *ACM National Conference* (1972), pp. 451–457.

[Par02] Paradiso A.: An algebra for combining mpeg-4 compliant facial animations. In *Online Proceedings; PRICAI International Workshop on Lifelike Animated Agents: Tools, Affective Functions, and Applications* (2002), Prendinger H., (Ed.). Tokyo, Japan, 19th of August 2002.

[PB03] Pelachaud C., Bilvi M.: Computational model of believable conversational agents. In *Communication in Multiagent Systems* (2003), Huget M.-P., (Ed.), Springer, pp. 300–317.

[PDR96] Planalp S., DeFrancisco V. L., Rutherford D.: Varieties of cues to emotion in naturally occurring situations. *Cognition and Emotion 10*, 2 (1996), 137–153.

[PDS*04] PAIVA A., DIAS J., SOBRAL D., AYLETT R., SOBREPEREZ P., WOODS S., ZOLL C., HALL L.: Caring for agents and agents that care: Building empathic relations with synthetic agents. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems* (2004), IEEE Computer Society, pp. 194–201.

[PDS*05] PAIVA A., DIAS J. A., SOBRAL D., AYLETT R., WOODS S., HALL L., ZOLL C.: Learn by feeling: Evoking empathy with synthetic characters. *Applied Artificial Intelligence 19*, 3–4 (2005), 235–266.

[Pet05] PETERS C.: Direction of attention and interest using gaze behaviour. In *Proceedings of the 5th International Working Conference on Intelligent Virtual Agents* (September 2005), Panayiotopoulos T., Gratch J., Aylett R., Ballin D., Olivier P.„ Rist T., (Eds.), vol. 3661 of *Lecture Notes in Computer Science*, Springer, pp. 215–228.

[PG96] PERLIN K., GOLDBERG A.: Improv: A system for scripting interactive actors in virtual worlds. In *Proceedings of SIGGRAPH* (1996), pp. 205–216.

[PI01] PRENDINGER H., ISHIZUKA M.: Let's talk! socially intelligent agents for language conversation training. In *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* (2001), vol. 31, pp. 465–471.

[PI02] PRENDINGER H., ISHIZUKA M.: Evolving social relationships with animate characters. In *Proceedings of the Symposium of the AISB Convention on Animating Expressive Characters for Social Interactions* (2002), pp. 73–78.

[Pic97] PICARD R. W.: *Affective computing*. MIT Press, 1997.

[PK27] PONDER E., KENNEDY W. P.: On the act of blinking. *Quarterly journal of experimental psychology 18* (1927), 89–110.

[PK02a] PIRKER H., KRENN B.: *Assessment of Markup Languages for Avatars, multimedia and multimodal system*. Tech. rep., ÖFAI, 2002.

[PK02b] PLANALP S., KNIE K.: *Integrating verbal and nonvrebal emotion(al) messages*. The verbal communication of emotions: interdisciplinary perspectives. Lawrence Erlbaum Associates, 2002, ch. 3, pp. 55–77.

[Plu80] PLUTCHIK R.: A general psychoevolutionary theory of emotion. In *Emotion: Theory, research, and experience*, Plutchik R., Kellerman H., (Eds.). Academic press, inc, 1980, pp. 3–33.

[PM03] PRASHANT C., MEYER J.: Modelling an infinite emotion space for expressionistic cartoon face animation. In *In Proceedings of the 6th IASTED International Conference on Computers, Graphics, and Imaging (CGIM2003)* (August 2003), pp. 13–18.

[PMI05] PRENDINGER H., MORI J., ISHIZUKA M.: Using human physiology to evaluate subtle expressivity of a virtual quizmaster in a mathematical game. *International journal of human-computer studies 62*, 2 (February 2005), 231–245.

[PP00] POGGI I., PELACHAUD C.: Emotional meaning and expression in performative faces. In *Affective Interactions: Towards a New Generation of Computer Interfaces*, Paiva A., (Ed.), vol. 1814 of *Lecture Notes in Computer Science*. Springer-Verlag, 2000, pp. 182–195.

[PP02a] PELACHAUD C., POGGI I.: Subtleties of facial expressions in embodied agents. *Journal of Visualization and Computer Animation (Special issue on Graphical Autonomous Virtual Humans) 13*, 5 (2002), 301–312.

[PP02b] POGGI I., PELACHAUD C.: *Signals and meanings of gaze in animated faces*. Language, Vision and Music: Selected papers from the 8th International Workshop on the Cognitive Science of Natural Language Processing. John Benjamins, 2002, pp. 133–Ű144.

[PP05] PRADA R., PAIVA A.: Intelligent virtual agents in collaborative scenarios. In *Proceedings of the 5th International Working Conference on Intelligent Virtual Agents* (September 2005), Panayiotopoulos T., Gratch J., Aylett R., Ballin D., Olivier P.„ Rist T., (Eds.), vol. 3661 of *Lecture Notes in Computer Science*, Springer, pp. 317–325.

[PPB*05] PETERS C., PELACHAUD C., BEVACQUA E., MANCINI M., POGGI I.: A model of attention and interest using gaze beahviour. In *Proceedings of the 5th International Working Conference on Intelligent Virtual Agents* (September 2005), Panayiotopoulos T., Gratch J., Aylett R., Ballin D., Olivier P.„ Rist T., (Eds.), vol. 3661 of *Lecture Notes in Computer Science*, Springer, pp. 229–239.

[PPdR*05] POGGI I., PELACHAUD C., DE ROSIS F., CAROFIGLIO V., DE CAROLIS B.: Greta. a believable embodied conversational agent. In *Multimodal Intelligent Information Presentation (Text, Speech and Language Technology Vol. 27)* (2005), Stock O., Zancanaro M., (Eds.), Kluwer.

[PPL86] PATTERSON M. L., POWELL J. L., LENIHAN M. G.: Touch, compliance and interpersonal affect. *Journal of Nonverbal Behaviour 10*, 1 (1986), 41–50.

[PPS01] PATERSON H. M., POLLICK F. E., SANFORD A. J.: The role of velocity in affect discrimination. In *J D Moore & K Stenning (Eds.) Twenty-Third annual conference of the cognitive science society* (August 2001), Lawrence ealrbaum associates, inc., pp. 756–761.

[PSB01] PERTAUB D.-P., SLATER M., BARKER C.: An experiment on fear of public speaking in virtual reality. In *Medicine meets Virtual Reality* (2001), pp. 372–378.

[PSB02] PERTAUB D.-P., SLATER M., BARKER C.: An experiment on public speaking anxiety in response to three different types of virtual audience. *Presence: Teleoperators and Virtual Environments 11*, 1 (February 2002), 68–78.

[PVP*03] PRADA R., VALA M., PAIVA A., HÖÖK K., BULLOCK A.: Fantasya - the duel of emotions. In *Proceedings of the 4th International Workshop on Intelligent Virtual Agents (Lecture Notes in Computer Science vol. 3661)* (September 2003), Rist T., Aylett R., Ballin D.„ Rickel J., (Eds.), vol. 2792, Springer-Verlag, pp. 62–66.

[PW96] PARKE F. I., WATERS K.: *Computer Facial Animation*. A K Peters Ltd., 1996.

[RA05a] REHM M., ANDRÉ E.: Catch me if you can: exploring lying agents in social settings. In *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems (AAMAS)* (2005), ACM Press, pp. 937–944.

[RA05b] REHM M., ANDRÉ E.: Informing the design of embodied conversational agents by analyzing multimodal politeness behaviors in human-human communication. In *Proceedings of the AISB Symposium for Conversational Informatics* (2005).

[RA05c] REHM M., ANDRÉ E.: Where do they look? gaze behaviors of multiple users interacting with an embodied conversational agent. In *Proceedings of the 5th International Workshop on Intelligent Virtual Agents (Lecture Notes in Computer Science)* (2005), Panayiotopoulos T., Gratch J., Aylett R., Ballin D., Olivier P.„ Rist T., (Eds.), vol. 3661, Springer, pp. 241–252.

[RCB98] ROSE C., COHEN M. F., BODENHEIMER B.: Verbs and adverbs: Multidimensional motion interpolation. *IEEE Computer Graphics and Applications 18*, 5 (1998), 32–40.

[RD05] RICHARDSON D. C., DALE R.: Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science: A Multidisciplinary Journal 29*, 6 (2005), 1045–1060.

[Rei97] REILLY S. N.: A methodology for building believable social agents. In *Proceedings of the first international conference on Autonomous agents* (1997), ACM Press, pp. 114–121.

[RFD97] RUSSELL J. A., FERNANDEZ-DOLS J. M.: *The psychology of facial expression*. Cambridge university press, 1997.

[RG91] RAO A. S., GEORGEFF M. P.: Modelling rational agents within a bdi-architecture. In *Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning* (1991), Morgan Kaufmann.

[RGBC96] ROSE C., GUENTER B., BODENHEIMER B., COHEN F. M.: Efficient generation of motion transitions using space-time constraints. In *Proceedings of SIGGRAPH* (August 1996).

[RGP93] ROSSBERG-GEMPTON I., POOLE G. D.: The effect of open and closed postures on pleasant and unpleasant emotions. *The Arts in Psychotherapy 20*, 1 (1993), 75–82.

[RKK03] RAOUZAIOU A., KARPOUZIS K., KOLLIAS S. D.: Online gaming and emotion representation. In *Proceedings of the International Workshop on Very Low Bitrate Video Coding (VLBV)* (2003), pp. 298–305.

[RL98] RICKEL J., LEWIS J. W.: Animated agents for procedural training in virtual reality: Perception, cognition and motor control. In *Applied Artificial Intelligence* (1998), vol. 13, pp. 343–382.

[RN96] REEVES B., NASS C.: *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge university press, 1996.

[RPP*03] ROSIS F. D., PELACHAUD C., POGGI I., CAROFIGLIO V., CAROLIS B. D.: From greta's mind to her face: modelling the dynamics of affective states in a conversational embodied agent. *International Journal of Human-Computer Studies 59*, 1-2 (July 2003), 81–118.

[RPPN02] RUTTKAY Z., PELACHAUD C., POGGI I., NOOT H.: Excercises of style for virtual humans. In *Proceedings of the Symposium of the AISB Convention on Animating Expressive Characters for Social Interactions* (2002), Canamero L., Aylett R., (Eds.), Advances in Consciousness Research, John Benjamins, pp. 73–78.

[RSH*05] ROMANO D. M., SHEPPARD G., HALL J., MILLER A., MA Z.: Basic: A believable, adaptable socially intelligent character for social presence. In *Proceedings of the 8th Annual International Workshop on Presence* (September 2005), Slater M., (Ed.).

[Rus97] RUSSELL J. A.: Reading emotion from and into faces: Resurrecting a dimensional-contextual perspective. In *The Psychology of Facial Expression*, Russel J. A., Fernández-Dols J. M., (Eds.). Cambridge University Press, 1997, pp. 295–320.

[Sch59] SCHACHTER S.: *The psychology of affiliation*. Stanford University Press, 1959.

[Sch64] SCHEFLEN A. E.: The significance of posture in communication systems. *Psychiatry 27* (1964), 316–331.

[Sch02] SCHROEDER R.: *The Social Life of Avatars; Presence and Interaction in Shared Virtual Environments*. Springer-Verlag New York, Inc., 2002.

[Sch04] SCHEUTZ M.: Useful roles of emotions in artificial agents: A case study from artificial life. In *Proceedings of AAAI* (2004), AAAI Press.

[SERS00] SCHIANO D. J., EHRLICH S. M., RAHARDJA

K., SHERIDAN K.: Face to interface: facial affect in (hu)man and machine. In *Proceedings of SIGCHI* (2000), pp. 193–200.

[SH95] SCHOUWSTRA S. J., HOOGSTRATEN J.: Head position and spinal position as determinants of perceived emotional state. *Perceptual and Motor Skills 81*, 2 (October 1995), 673–674.

[Slo01a] SLOMAN A.: Beyond shallow models of emotions. In *Cognitive Processing* (2001), vol. 2, pp. 177–198.

[Slo01b] SLOMAN A.: Varieties of affect and the cogaff architecture schema. In *Proceedings of the Symposium on Emotion, Cognition, and Affective Computing - AISB Convention* (2001), pp. 39–48.

[Som69] SOMMER R.: *Personal space: The behavioural basis of design*. Prentice-Hall, 1969.

[SPBC04] SLATER M., PERTAUB D.-P., BARKER C., CLARK D.: An experimental study on fear of public speaking using a virtual environment. In *Proceedings of the third International Workshop on Virtual Rehabilitation* (September 2004).

[SPS99] SLATER M., PERTAUB D.-P., STEED A.: Public speaking in virtual reality: Facing an audience of avatars. In *IEEE Computer Graphics and Applications* (March 1999), vol. 19, pp. 6–9.

[SS62] SCHACHTER S., SINGER J. E.: Cognitive, social and physiological determinants of emotional states. *Psychological Review 69*, 5 (September 1962), 379–399.

[SS72] SCHEFLEN A. E., SCHEFLEN A.: *Body language and social order: communication as behavioral control*. Prentice-Hall inc., 1972.

[SS97] SMITH C. A., SCOTT H. S.: A componential approach to the meaning of facial expressions. In *The Psychology of Facial Expression*, Russel J. A., Fernández-Dols J. M., (Eds.). Cambridge University Press, 1997, pp. 295–320.

[SS01] SAKAGUCHI H., SAKAKIBARA M.: Final fantasy: The spirits within. 16mm 106 min, Square Pictures, 2001.

[SS02] SLATER M., STEED A.: *Meeting people virtually: experiments in shared virtual environments*. The social life of avatars: presence and interaction in shared virtual environments. Springer-Verlag New York, Inc., 2002, ch. 9, pp. 146–171.

[SSL00] SCHEUTZ M., SLOMAN A., LOGAN B.: Emotional states and realistic agent behaviour. In *Proceedings GAME-ON* (November 2000).

[SVS05] SANCHEZ-VIVES M. V., SLATER M.: From presence to consciousness through virtual reality. *Nature Reviews Neuroscience 6*, 4 (2005), 332–339.

[TBS*98] TROMP J. G., BULLOCK A., STEED A., SADAGIC A., SLATER M., FRÉCON E.: Small group behaviour experiments in the coven project. In *IEEE Computer Graphics and Applications* (1998), vol. 18, pp. 53–63.

[TC99] THÓRISSON K. R., CASSELL J.: The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents. In *Applied Artificial Intelligence* (May 1999), vol. 13, pp. 519–538.

[TCP97] TORRES O. E., CASSELL J., PREVOST S.: Modeling gaze behavior as a function of discourse structure. In *First International Workshop on Human Computer Conversations* (1997).

[TF79] TEASDALE J. D., FOGARTY S. J.: Differential effects of induced mood on retrieval of pleasant and unpleasant events from episodic memory. *Journal of abnormal psychology 88*, 3 (June 1979), 248–257.

[TF85] THORSON E., FRIESTAD M.: *The effects on emotion on episodic memory for television commercials*. Advances in consumer psychology. Lexington, 1985, pp. 131–136.

[Thó97] THÓRISSON K. R.: Layered modular action control for communicative humanoids. In *Proceedings of Computer Animation* (1997), IEEE Computer Society, pp. 134–143.

[Thó99] THÓRISSON K. R.: Mind model for multimodal communicative creatures and humanoids. In *International Journal of Applied Artificial Intelligence* (1999), vol. 13, pp. 449–486.

[Thó02] THÓRISSON K. R.: Natural turn-taking needs no manual: Computational theory and model, from perception to action. In *Multimodality in Language and Speech Systems*, Granström B., House D.„ Karlsson I., (Eds.). Kluwer Academic Publishers, 2002.

[Tha96] THAYER R. E.: *The origin of everyday moods*. Oxford University Press, 1996.

[Tho96] THORISSON KRISTINN R.: Toonface: A system for creating and animating interactive cartoon faces. In *Technique report, MIT Media Laboratory, Learning and Common Sense* (April 1996), pp. 13–18.

[TI65] TOMKINS S. S., IZARD C. E.: *Affect, cognition and personality*. Springer publishing company, inc., 1965.

[TJ81] THOMAS F., JOHNSON O.: *Disney Animation - The illusion of life*. Abbeville Press, 1981.

[TLPD05] THÓRISSON K. R., LIST T., PENNOCK C., DIPIRRO J.: Whiteboards: Scheduling blackboards for semantic routing of messages & streams. In *AAAI Workshop on Modular Construction of Human-Like Intelligence* (2005), Thórisson K. R., Vilhjálmsson H.„ Marsella S., (Eds.), pp. 8–15.

[Tom80] TOMKINS S. S.: *Affect as Amplification: some* *modification in theory*. Academic press, inc, 1980, pp. 3–33.

[Tom04] Tomoko K.: Interpretation of emotionally expressive characters in an intercultural communication. *Proceedings of the 8th International Conference Knowledge-Based Intelligent Information and Engineering Systems (Lecture Notes in Computer Science) 3214* (2004), 862–868.

[UHN98] Ushida H., Hirayama Y., Nakajima H.: Emotion model for life-like agent and its evaluation. In *AAAI/IAAI: Proceedings of the fifteenth national/tenth conference on Artificial intelligence/Innovative applications of artificial intelligence* (1998), American Association for Artificial Intelligence, pp. 62–69.

[VBG*04] Vinayagamoorthy V., Brogni A., Gillies M., Slater M., Steed A.: An investigation of presence response across variations in visual realism. In *The 7th Annual International Presence Workshop* (October 2004), pp. 148–155.

[VC98] Vilhjálmsson H. H., Cassell J.: Bodychat: Autonomous communicative behaviours in avatars. In *International Conference on Autonomous Agents and multi-agent systems* (May 1998), vol. 2, pp. 269–276.

[Vel97] Velásquez J. D.: Modelling emotions and other motivations in synthetic agents. In *American Association for Artificial Intelligence* (1997), pp. 10–15.

[Vel98] Velásquez J. D.: A computational model of emotional conditioning in the brain. In *Grounding Emotions in Adaptive Systems at the 5th International Conference of the Society for Adaptive Behaviour (SAB)* (August 1998).

[VGSS04] Vinayagamoorthy V., Garau M., Steed A., Slater M.: An eye gaze model for dyadic interaction in an immersive virtual environment: Practice and experience. *Computer Graphics Forum 23*, 1 (March 2004), 1–11.

[Vil05] Vilhjálmsson H. H.: Augmenting online conversation through automated discourse tagging. In *Proceedings of the 38th Hawaii International Conference on System Sciences (HICSS)* (2005), IEEE Computer Society.

[VSS05] Vinayagamoorthy V., Steed A., Slater M.: Building characters: Lessons drawn from virtual environments. In *Towards Social Mechanisms of Android Science: A CogSci 2005 - XXVII Annual Conference of the Cognitive Science Society workshop* (July 2005), pp. 119–126.

[VSS06] Vinayagamoorthy V., Steed A., Slater M.: The influence of a virtual character posture model on reported presence in an immersive virtual environment. In *ACM SIGGRAPH International Conference on Virtual Reality Continuum and Its Applications - VRCIA* (June 2006), pp. 229–236.

[Wal98] Wallbott H. G.: Bodily expression of emotion. *European Journal of Social Psychology 28*, 6 (1998), 879–896.

[WBJ68] Watzlawick P., Beavin J. H., Jackson D. D.: *Pragmatics of human communication*. Faber & Faber, 1968.

[WDRG72] Wiener M., Devoe S., Robinson S., Geller J.: Nonverbal behaviour and nonverbal communication. *Psychological Review 79*, 3 (May 1972), 185–214.

[WKSS00] Wehrle T., Kaiser S., Schmidt S., Scherer K. R.: Studying the dynamics of emotional expression using synthesized facial muscle movements. *Journal of Personality and Social Psychology 78*, 1 (January 2000), 105–119.

[WW86] Walters K. L., Walk R. D.: Perception of emotion from body posture. *Bulletin of the Psychonomic Society 24*, 5 (1986), 329–329.

[WW88] Walters K. L., Walk R. D.: Perception of emotion from moving body cues in photographs. *Bulletin of the Psychonomic Society 26*, 2 (1988), 112–114.