

# Requirements of non-verbal communication in believable synthetic agents

John Shearer

Informatics Research Institute  
University of Newcastle upon Tyne  
Newcastle upon Tyne  
NE1 7RU  
United Kingdom  
John.Shearer@ncl.ac.uk

Philip Heslop

Informatics Research Institute  
University of Newcastle upon Tyne  
Newcastle upon Tyne  
NE1 7RU  
United Kingdom  
Philip.Heslop@ncl.ac.uk

Patrick Olivier

Informatics Research Institute  
University of Newcastle upon Tyne  
Newcastle upon Tyne  
NE1 7RU  
United Kingdom  
P.L.Olivier@ncl.ac.uk

Marco De Boni

Unilever Corporate Research  
Unilever R&D Colworth  
Sharnbrook, Bedford  
MK44 1LQ  
United Kingdom  
Marco.De-Boni@unilever.com

## Abstract

Realism for synthetic characters, both in computer games and conversational agent mediated applications, requires both visual and behavioural fidelity. One significant area of synthetic character behaviour, that has to date received little attention, is non-verbal communication. In identifying the scope and participants of non-verbal communication in computer games we first review the range of spatial and task scenarios that are relevant. We then select four principal categories of non-verbal behaviour: gesture; gaze; proxemics; self-adaptors; and both summarise their role in communication and propose their incorporation in the design of non-player characters. In conclusion we review the question of how to capture the non-verbal behaviour of players or provide them with interaction techniques that might facilitate non-verbal communication with players and non-player characters alike.

## 1 Introduction

Perhaps the most significant current challenge for synthetic agents in games is the integration of narrative and *intelligent* character behaviour with a view to further enhancing user engagement. The development of non-player characters has primarily focussed on high level coordination of non-player character behaviours (typically through the characterisation of state space) and little attention has been paid to low level, non-physical, interactions between characters (both non-player characters and player characters) such as different forms of verbal and non-verbal communication.

Realistic interaction between characters is a challenging problem. In simple terms it requires non-player characters to have sophistication in their beliefs, desires, and intentions equivalent to the expectations that a player might have of a real equivalent character (within the narrow domain of the game). Although, one should keep in mind the possibility of an ‘uncanny valley’ (Mori, 1970) –

that as characters become more humanlike, with a positive emotional response from a human, there becomes a point where the response suddenly becomes strongly negative – the character is almost human, but is a very strange human – something is perceived as being not quite right. This is clearly contrary to the desired effect.

Crucially, non-player characters must also have beliefs as to the beliefs, desire and intentions of other characters, and therefore require some ability to monitor the activity of other characters for intentional behaviour. This then allows for communication (the intentional influence of other’s beliefs).

Of course, this requires a non-player character to distinguish between behaviour that is, and is not, intentionally influencing another’s beliefs, as many behaviours are interpreted differently depending on the context. For example, opening a door can be an action (enabling continuation of a route) and communication (showing another person respect). Humans (and many higher mammals) are extremely

good at separating basic actions from actions intended to communicate (as a conscious act or otherwise), and, in the course of social interactions, people constantly analyse each other's behaviour (e.g. speaking, moving, gesturing, touching) to understand their (the other's) beliefs, desires and intentions. It has been shown that people also do the same for artificial entities (Reeves and Nass, 1996). That is, people have a tendency to attribute intention to all things they interact with, even if those objects have no possible mechanism from which intention might derive (according to our present scientific understanding of the world), for example, by shouting at a household object if it is not working correctly. Inevitably, players ascribe intention (often very complex intention) to non-player characters even when, their behaviour, and/or the underlying control architecture, is very limited.

Non-player characters in present-day games often incorporate a restricted notion of desire and intention, but have no significant ability to perceive the desires and intentions of other characters (and thereby infer their beliefs and other aspects of internal state). Thus, intention is usually exhibited through direct action – in a first person shooter non-player characters display their intention to frag (kill) a character simply by shooting at that character. Indeed, in games where the role of non-player characters is not so clearly defined, players quickly habituate to repetitive behaviour as this “supporting cast” mills around according to simple rules or scripts.

A discussion of the development of interactions in all areas of games is beyond the scope of our discussion. Instead, we address communication between characters, specifically non-verbal communication through gesture, facial expression, eye gaze and other bodily movement. That is, we address non-verbal interactions between characters that are not direct actions (i.e. not actions such as attacking). Most games portray the intentionality and personality of characters with cut-scenes – scenes inserted, pausing game-play, at various points of the game – with fully animated characters using non-verbal and verbal behaviour in an attempt to provide an engaging narrative for a game. In contrast, during actual game-play a character's behaviour is much simpler, and very little intentionality is usually displayed.

The advent of high performance 3D graphics hardware and software, and high quality audio has given rise to the use of geometrically detailed, skeletally animated characters, that exhibit some non-verbal behaviour. For example, *Half-Life 2* (Valve, 2004) makes extensive use of non-verbal behaviour in non-player characters, to the extent that player engagement and narrative development can occur effectively without the use of cut-scenes. Non-player characters attempt to talk with the player; if engaged in talking to a player character

they will: move around to maintain eye contact during the conversation; respond to the player's actions; and show expressions and gestures appropriate to the game-play, thus attempting to engage the player with the narrative.

While the characters within *Half-Life 2* appear to have desires, personality and intentions, their actual behaviour is, in fact, highly scripted. Most non-player characters have only a relatively small set of simple behaviours and a number of complex scripts that only occur once or twice. These complex behaviours replace the traditional cut-scene, but without the need to break game-play. However, the integrity of the non-verbal behaviour relies on the skill of the artists in coordinating the non-player character's action with the game-play and the non-player character's speech. This can clearly be seen when observing non-player characters' interactions with each other. Aside from pre-scripted cut scenes, non-player characters do not interact with each other, other than to avoid collision. Even then, they still walk on paths that lead to collision, but evade each other at the last moment (sometimes accompanied with a sound effect along the lines of ‘watch where you are going’). Indeed, following an individual non-player character through the environment reveals that they are predominantly on pre-fixed paths that simply enter the environment at one point and walk off at another. Some of the inhabitants even walk in a constant looping path around the environment.

Similar behaviour is exhibited in other games, such as *Grand Theft Auto 3* (Rockstar Games, 2001). Although more densely populated, the non-player characters exhibit similarly aimless behaviour, reacting only to certain events (such as explosions or being attacked) but otherwise wandering through the environment with no apparent destination.

There are games in which the behaviour of non-player characters is more sophisticated, such as ‘*The Sims*’ (Maxis, 2000), where the characters fall half-way between being user controlled and AI controlled. The characters converse with each other and utilise objects in their environment. Though, even in this case the characters exhibit only the desire to satiate certain undesirable internal states, such as hunger and tiredness, by using the objects and characters around them. They have no longer term and more complex goals, and the conversations between characters do not actually convey any specific information but rather the act of conversation itself satisfies some current short-term need; the characters can obtain the same effect by using an object in the environment, such as reading a book or listening to the radio.

In practice the scope of application of non-verbal behaviour extends far beyond the replacement of the cut-scene and in the next section we consider the applicability to games of the relevant spatial and

task contexts over which non-verbal communication is known to be utilised in humans.

## 2 Spatial and task context

Communication can be considered to occur in four different task contexts: cooperation; coaction; competition; conversation (Knapp and Daly, 2002, p177). In other words, communication occurs in order for some number of parties to: perform a task together (cooperation), to exist in the same vicinity (coaction), to perform a task at the expense of another (competition), and to entertain or pass on information (conversation). Communication varies across these different contexts and also with the physical proximity of the communicating parties. As already described, non-verbal communication provides information as to the beliefs, desires, and intentions of a character, or alternatively it can be considered as providing indicators as to another character's: cognitive state; emotional state; physical state; intentional state; attentional state; perceptual state; interactional state; social state.



Figure 1a: Cooperation in intimate space.

Figure 1b: Coaction in social space.



Figure 1c: Conversation in personal space.

Figure 1d: Competition in personal space.



Figure 1e: Cooperation at a distance.

Figure 1f: Conversation in social space.

Figure 1: Examples from Half-Life 2 of spatial and task contexts for non-verbal communication

A set of distinct but common communication situations (spatial and task contexts) for computer games is illustrated in Figure 1 using screenshots from Half-Life 2. Figure 2 maps out the range of spatial and task context for these examples. Synthesizing non-verbal communicative behaviour for conversational partners at close physical proximity is particularly difficult, due to the full movement of a character (both body and face) being

visible in detail to the player. Furthermore, people are highly attuned to interactions in intimate, personal, and social spaces and are sensitive to many subtle cues and nuances in non-verbal behaviour. Thus, at close proximity players are very aware of errors, un-realistic, or unnatural behaviour in non-player character. At further distances less detail of a character's movement is apparent. Moreover, there is a significant transition in non-verbal communication from situations where intimate verbal communication is possible to those where it is not. The sensitivity of non-verbal behaviour to proximity is due to a number of factors, including the more public nature of non-verbal gesture in open spaces, and the requirement on particular physical behaviour to carry the full communicational load (e.g. subtleties in gaze and facial expression are not visible at a distance).

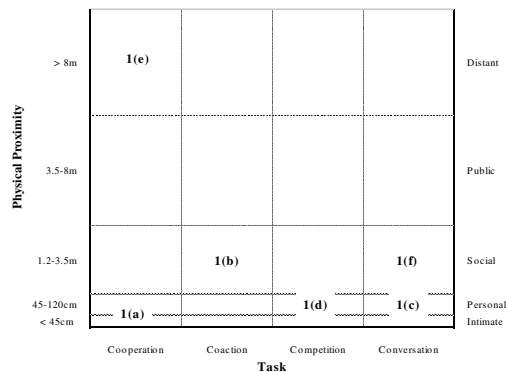


Figure 2: Task and spatial contexts.

Figure 1a shows an example of cooperation in intimate space. The male character demonstrates his attentional state – that he is attending to the female character – with his body orientation, face orientation, and gaze direction. Of course, people are rarely static, but different non-verbal channels (e.g. face orientation, body orientation, gaze direction, body position) are closely coordinated in demonstrating attention. Thus, the male character could look away but still communicates his attention sufficiently with the rest of his body. It should be noted that between unfamiliar participants strong or constant facing or looking at a person is widely considered an aggressive signal, or is considered rude, or at least off-putting, in an interaction (Knapp and Daly, 2002, p380). Figure 1a also illustrates non-verbal communication using facial expressions and kinaesthetic (touching) behaviour.

Figure 1e illustrates a situation at the other end of the spatial scale, cooperation at a distance between the player and a non-player character (in fact, navigation and negotiation, a subset of cooperation). The non-player character shown and

the player will collide if they do not arrive at an agreement as to how to pass one another and communicate this – the characters must cooperate through the use of non-verbal communication to resolve a potential conflict. In the real world, people in this situation use a range of subtle non-verbal mechanisms such as gaze and body turning to initiate and mutually negotiate space. Non-player characters in Half-Life 2 will avoid the player, but will not exhibit non-verbal communication in doing so and simply move around the players as they approach. It would clearly be more realistic for the non-player characters to use non-verbal communication as humans do to negotiate past the player. Without this level of non-verbal behaviour it is difficult for the player to decide which way to move out of the way (indeed they do not need to) and it might be argued that this absence of social convention (and the ability to break them) both undermines the engagement of players with the game and limits their expressivity. For example, walking into another character and the social conflict that arises may serve to further or undermine a player's ends.

Between the proximal and distant spatial scales are social spaces and figure 1f is an example of conversation in a social space. Here non-verbal behaviour facilitates a number of aspects of the interaction (and the dialog in particular) including the mediation of conversation flow, such as whose turn it is to speak (interactional state). Turn-taking mediation is a complex coordination of behaviours, but in simple terms speakers provide turn-swap opportunities (such as, a slightly prolonged pause, or a look up into the eyes), at which point other listeners *can*, if they choose, take a turn. If not, then the speaker will continue. Additionally, the non-listener can indicate that they would like to speak, with signals such as: increased eye contact; leaning forward; standing taller (Duncan and Fiske, 1977). Turn-taking mediation is not required in Half-Life 2 because the game developers have not allowed the player to speak, but it is potentially a very important component of games that hope to include natural language interactions (particularly spoken interaction) between player and non-player characters.

Finally, figures 1b, 1c and 1d illustrate the remaining task contexts: coaction, conversation and cooperation, and competition. Characters sharing the same approximate area of space engage in coaction behaviour, corresponding to mutual monitoring – this can be interpreted as communication by virtue of the fact that watching a character implies that you might react to it – that is, there is an implied reason (intention) for watching. Coaction can be considered the default task context, which develops into the other contexts. Competition contexts give rise to distinctly different forms of non-verbal communication from other contexts, but these still

serve to communicate internal states. In figure 1d the raised baton serves to communicate “you have crossed a line – back off or I will hit you”.

### 3 Non-verbal communication

Non-verbal communication in non-player characters should be based on the non-verbal communication that people actually use and exhibit in human-human interactions. Hence, the forms of non-verbal communication that are particularly relevant to the synthesis of non-verbal communication in virtual characters are: proxemics (the use and arrangement of the self in the physical world); gaze (where the eyes are looking and pupil dilation); gesture (movement of hands and head); self-adaptors (movements that serve to alter the self); passive communication (communication that occurs without a specific action).

It should be noted that there is a very significant variation among individuals in the precise nature of their non-verbal communication. Non-verbal communication is influenced by age; gender; social class/hierarchy position; culture; presence of others; illness; physical ability, as well as the spatial and task context. This presents problems for both the study and the synthesis of non-verbal behaviour. Fidelity in non-verbal communication requires that each distinct non-player character employs its own characteristic behaviour in this respect. In the following sections we characterise non-verbal communication in people, generally from a Western/Anglo-Saxon cultural perspective, with a view to identifying which aspects should be addressed in the design of engaging virtual characters for computer games.

#### 3.1 Proxemics

Proxemic communication is defined as the closeness and arrangement of the self in physical world compared to others (Hall, 1966). This is the use of personal space and territory, and although it is difficult to get an accurate sense of distance in games (people rarely feel their personal space is invaded in games), the use of space does have an impact on player behaviour. Moreover, as games become more immersive (due to improved graphics, larger screens, head-mounted displays, immersive environments, and surround sound) and dialog with non-player characters becomes a common occurrence, the impact of proxemics will increase. As discussed, non-verbal communication varies significantly with distance (and spatial arrangement), so its consideration is therefore important in the synthesis of non-player character behaviour.

Proxemics also exhibit some of the largest (but consistent) cultural variations, especially in conversation. In order of increasing distance (or decreasing proximity) the spaces for UK/US culture

(Hall, 1966) are: *intimate space* for embracing, touching or whispering (15-45 cm, 6-18 inches); personal space for interactions among good friends (45-120 cm, 1.5-4 feet); social space for interactions among acquaintances (1.2-3.5 metres, 4-12 feet); public space used for public speaking (over 3.5 metres, 12 feet). Additionally, we can define *distant space* for interaction at over 8 metres (25 feet), where people still interact but do not conduct conversations.

Proxemic behaviour stems from the idea of territory. Territories vary dynamically and are dependent upon many different factors, but the categories above are likely to be sufficient for basic synthesis. The requirement for synthesis is that non-player characters use proxemics as people do – to be socially correct use: social space for interacting with relatively unknown people; personal space for people known well; intimate space with people known very well. Violation of these rules is socially incorrect. Invasion of personal space can be intimidating, flirtatious, or could be the result of interaction between participants from different cultural backgrounds. Whilst the requirements of a task, such as needing to be close to attend to a wound, may override the usual norms, in such circumstances the progression through the spaces is still mediated by non-verbal behaviour.

The synthesis of complex proxemic behaviour requires further parameterisation with respect to types of territory: primary; secondary; public (Altman, 1975). Primary territories are the exclusive domain of the owner (such as a home), secondary territories are those felt to be partly owned (such as the local pub), and public territories those available to almost anyone for temporary ownership (such as a park bench). The idea of temporary ownership, is important as the temporary ownership means a person behaves very differently towards that object or space while they ‘own’ it (see Knapp and Daly, 2002).

### 3.2 Gaze and eye-based communication

The use of the eyes is an important component of human-human communication, and it involves far more than just what a person is looking at. Peoples’ eyes are constantly moving from place to place (not smoothly, but jumping from one locus of attention to another). As Knapp and Daly note, “we associate various eye movements with a wide range of human expressions: downward glances are associated with modesty; wide eyes with frankness, wonder, naïveté, or terror” (2002, p370). Where the eyes are looking (or gaze) is the primary (and most obvious) form of communication by gaze, but people are also sensitive to, and react to, pupil dilation/constriction. There is also a close relationship between gaze

behaviour and facial expression (not considered here).

Kendon (1967) identifies four functions of gaze behaviour (in addition to looking at specific items for information gathering), and Knapp (2002) builds on this, classifying five functions of gaze as:

1. regulating the flow of communication;
2. monitoring feedback;
3. reflecting cognitive activity;
4. expressing emotions;
5. communicating the nature of an interpersonal relationship [added by Knapp]

Within the context of a non-player character, the five categories pose challenges for synthesis that are considerably more challenging than simply having the character look at where it is interested. The regulation of communication flow, gazing briefly at another person (specifically at the face) establishes an obligation to interact; further and longer gazing shows a desire to increase the level of interaction; while decreased and shorter gazing desires a decrease in the level of interaction. During an interaction eye glances serve as turn-taking signals and also highlight grammatical breaks, conceptual unit breaks, and the ends of utterances (a sequence of speech separated from another by a marked gap). These glances also allow feedback on the interaction by monitoring the reactions of the other person.

When under increased cognitive load (trying to process difficult or complex ideas) both listeners and speakers tend to look away. It is thought that this averted gaze reflects a shift in attention from the external to the internal. Additionally, there is evidence that the eye gaze direction under this condition changes with different forms of cognitive load, and that it is linked to the active hemisphere of the brain (Ehrlichman and Weinberger, 1978; Weisz and Adam, 1993; Wilbur and Roberts-Wilbur, 1985).

The eyes are also a site for the display of emotions: surprise; fear; disgust; anger; happiness; sadness, in addition to blends of these and more complex emotions. Interestingly, in certain contexts emotions displayed with the eyes will not always match the facial emotional expressions (e.g. during emotional masking) and can be very transitory. However, people are adept at detecting emotional state from the eyes, and there is evidence that different emotions are in fact detected from different areas around the face (Ekman et al., 1971; Ekman and Friesen, 1975), but that it is the facial area *around* the eyes that displays the emotion, not the eyes themselves.

Finally, eye gaze can communicate the nature of an interpersonal relationship. Gazing and mutual gazing is found most in conversations with a high-status addressee, lower with a very high status addressee, and minimal with a very low-status addressee (Hearn, 1957; Efran, 1968). And, of

course, lovers have extended periods of mutual gazing (this also occurs with mothers and babies).

### 3.3 Gesture

Gesture is body movement that serves to communicate, mainly involving movement of the hands and the head. There are many forms of gesture, but of particular relevance for non-verbal communication synthesis in non-player characters are emblematic gesture (gesture with specific meanings that occur without speech) and spontaneous gesture (hand and head movements that occur with speech). Emblematic gestures are well defined in both their form and meaning and are therefore readily synthesised with standard skeletal animation and scheduling frameworks. For example, the 'come here' gesture is performed using the moving of a finger, fingers, hand, hands, arm, or arms towards the body from the direction of the addressee (often in a repetitive form). Synthesis of spontaneous gesture for non-player characters is significantly more challenging.

Spontaneous gesture is performed by people while speaking, in synchrony with the speech and is generally made with the head or hands. When needs be people are apt to use any available body part, or even the whole body (e.g. pointing with a foot when one's hands are full). Spontaneous gesture continues to be studied in depth by the psychology and psycholinguistics communities, and though most studies are descriptive in nature, recent research on growth points is leading to theories of how gesture relates to other cognitive processes (McNeill Lab, University of Chicago, 2006; McNeill, 2005).

It has been found that the gesture stroke (the semantic, or meaningful part) of the gesture commonly coincides with the *peak phonological stress* – the most emphasised phoneme – of the speech stream. Gesture is tied closely to the underlying speech and both speech and gesture are widely believed to be generated from a single underlying conception (McNeill, 1992). Indeed, spontaneous gesture can be complementary, supplementary, or contrastive to the speech. In other words, gesture can re-iterate or emphasise the speech, add information to the speech, or communicate something contradictory (or slightly different) from the associated speech. In contrast to speech, gesture has few constraints on how it is constructed. As McNeill notes (1992, p1) "the important thing about gestures is that they are *not* [original emphasis] fixed. They are free and reveal the idiosyncratic imagery of thought." McNeill identifies five categories of spontaneous gesture:

1. emphatic;
2. deictic;
3. cohesive;
4. metaphoric;
5. iconic.

Emphatic gesture (also known as beat gesture, or baton gesture) provides emphasis to: parts of speech; phrases; words; phonemes. Emphatic gesture consists of just two movement phases (up/down, in/out, left/right) with the transition from one phase to the other (such as, up to down) being the point of emphasis. These have little variation in form other than the scale and speed of the phase transition, with larger, faster transitions meaning more emphasis (within an individual). Emphatic gesture can, and frequently does, utilise all body parts, especially the head and hands, but additional movement of more of the body provides further emphasis. This form of spontaneous gesture is distinct from other forms in that it can overlay any other gesture as it indexes a part of speech rather than providing semantic content (though it is also commonly used independently).

Deictic gestures are simply pointing actions that refer to an object or objects, generally using fingers, hands, or head. This is complicated by the fact that deictic gesture can, in addition to referencing concrete objects, reference more abstract objects, such as where an object was previously, the physical space referred to previously with the introduction of an idea or object, or almost any abstract space, including time. For example when describing a cartoon involving two characters people will often reference one specific area of their gesture space for one character and another separate area for the other (McNeill, 1992). It should be noted that deictic gestures can take the form of whole body movement towards a space as well as the orientation of body parts.

Cohesive gestures serve to connect related parts of discourse that are temporally separated. For example, when listing items people often provide an emphatic gesture on each item. The emphatic gesture clearly marks each item, while the repetition of the same gesture form connects them together to say 'here's *one*, and *another*, and *another*, and *another* ...'. Cohesive gesture usually consists of the repetition of a specific gesture and so they require the use of other gesture forms (which could simply be emphatic gesture). Finally, metaphoric and iconic gestures are in essence pictorial. An iconic gesture is a pictorial animation of a concrete entity or action, while metaphoric gesture animates an abstract concept as if it were a physical object. For example, a speaker saying 'and he bends it way back' while illustrating the action of bending a stick is performing an iconic gesture (McNeill, 1992, p12). In contrast metaphoric gesture occurs in situations such as when a speaker says 'I had this great idea' and marks the 'great idea' with a cupped hand gesture (i.e. information container metaphor). Metaphoric gesture takes abstract ideas and grounds them in real entities. In practice only a few types of object are portrayed, mostly containing or enclosed

objects, though it is often not clear what object is being portrayed – it is just *some* object.

Iconic gestures are probably the most sophisticated and developed class of gesture; each gesture attempts to portray some aspect(s) of a situation, event, or object in the physical world and therefore absorb much of the complexity of the physical world. Typically the gesture portrays the most important and semantically salient features in the context of the interaction, and therefore a gesture about a specific object, such as a teacup, can vary significantly depending on the context. For example, a person talking about a teacup in context of drinking tea may perform a gesture of lifting a teacup by its handle, while if the context was about how much tea was in the cup, the gesture would be distinctly different – perhaps portraying the size of the teacup as opposed to how it is lifted. Iconic gesture is the form of gesture that most allows for the addition of extra information in the gestural channel and almost anything is permitted. For objects or ideas that are common in interaction the gesture forms portraying them can often become stylised, and eventually can become symbols, akin to emblems.

In addition to gesture accompanying speech, gesture can be used in a similar way to eye gaze and many vocal signals to regulate communication flow and the rhythm of interaction (Knapp and Daly, 2002, p252-3). Head nods are the most frequent form of these gestures, but hand and body gestures can also serve for flow regulation. Unsurprisingly, flow regulation gesture frequently coincides with flow regulation signals in other channels.

### 3.4 Self-adaptors

Self-adaptors are movements that serve to change the self, such as scratching an arm; as such they are typically not intended to serve a communicative function – their purpose is to adjust something about the self, but people read meaning into self-adaptors and so it can be considered a channel for communication. Self-adaptors can take on many forms and are distinctly idiosyncratic. The meanings assigned to them by conversational participants and bystanders vary considerably. For example, flicking hair out of the face, while a practical movement, is often interpreted as a flirtatious behaviour.

### 3.5 Passive Communication

Finally, another consideration to be taken into account is passive communication – communication that occurs without a specific action on the part of the communicator, and usually takes place at the co-action level of communication, and at a greater distance than conversational communication.

The communicator expresses information about themselves simply by their physical appearance, demeanour and observed behaviour. These factors

can be considered by the user (or indeed by other non-player characters) before initiating a more direct method of communication. For example, a large aggressive character might be approached with more caution than a small timid looking character. Similarly, witnessing a character commit some violent act would instil more caution in the user than a character who has been doing something less impactful.

Passive communication is often unintentional at the time of communication itself, but can be premeditated, for example wearing a certain type of clothing might convey a specific message.

## 4 Facilitating interaction

The discussion of non-verbal communication above concentrates on the different forms of non-verbal communication in an attempt to establish requirements for next generation virtual characters. Mechanisms for the synthesis of such behaviour are inevitably beyond the scope of this analysis but there have been a number of recent attempts to synthesise spontaneous gesture (Kopp and Wachsmuth, 2004; Olivier, 2003; Cassell, Vilhjálmsón, and Bickmore, 2001) and gaze behaviour. However, whilst we can imagine a situation whereby non-player characters have sophisticated cognitive models and the ability to both synthesise and interpret non-verbal behaviour, there is no mechanism for the player to communicate non-verbally with either the non-player characters or other players.

Present-day technology allows collection of full data on all aspects of human physical behaviour that may be a channel for non-verbal communication. This includes body position, body movement, hand shape, eye gaze direction, pupil dilation, facial expression, vocal behaviour, voice, and a variety of other biometrics. Unfortunately, such data can only be collected accurately using specialised (expensive) invasive equipment. With that commercial consideration, in addition to joystick (or controller), keyboard and mouse inputs it is only reasonable to expect relatively basic additional input devices for computer games in the near future. These devices include webcams and microphones, but also less standard input methods, such as ‘Dance mats’, light guns, and low point motion capture devices, such as the Gametrack 3D motion tracker.

Within current commercial and technological constraints, three alternatives for the control of player character non-verbal behaviour generation can be identified:

1. simulation – non-verbal communication of human players is simulated as for non-player characters, and is independent of the actual non-verbal communication of the human player.

2. augmentation – as for simulation, but specific controls are given to allow a player to explicitly alter the simulation, such as a slider to indicate how happy the player is, or a button to increase the level of interest in an interaction.
3. tracking and mapping – the human player is tracked using equipment such as webcams and the coarse features that can be identified are either mapped to explicit controls (as in the case of augmentation) or directly to the character animation.

The challenge for augmentation is to design an interface that is intuitive, non-obtrusive, and useful all at the same time. While the challenges in player non-verbal communication data collection are daunting, the challenges of understanding or recognising behaviours or meaning from that data are even more so. Gesture recognition, for example, is in its infancy and mostly addresses the use of gesture as an explicit interaction technique and little research has been conducted into the recognition or understanding of spontaneous gesture.

## References

- I. Altman. *The environment and social behaviour*. Brooks/Cole, Monterey, CA, 1975.
- J. Cassell, H. Vilhjálmsón, T. Bickmore. BEAT: the behavior expression animation toolkit. *Proceedings of SIGGRAPH '01*:477-486, Los Angeles, CA, 2001.
- S. Duncan and D.W. Fiske. *Face-to-face interaction: research, methods, and theory*. L. Erlbaum Associates, Hillsdale, N.J., 1977.
- J.S. Efran. Looking for approval: Effects on visual behaviour of approbation from person differing in importance. *Journal of Personality and Social Psychology*, 10:21-25, 1968.
- H. Ehrlichman and A. Weinberger. Lateral eye movements and hemispheric asymmetry: A critical review. *Psychological Bulletin*, 85:1080-1101, 1978.
- P. Ekman and W.V. Friesen. *Unmasking the face*. Prentice-Hall, Englewood Cliffs, NJ, 1975.
- P. Ekman, W.V. Friesen and S.S. Tomkins. Facial affect scoring technique: A first validity study. *Semiotica*, 3:337-58, 1971.
- S. Kopp and I. Wachsmuth. Synthesizing Multimodal Utterances for Conversational Agents. *The Journal of Computer Animation and Virtual Worlds*, 15(1):39-52, 2004.
- E.T. Hall. *The hidden dimension*. Doubleday, Garden City, N.Y., 1966.
- G. Hearn. Leadership and the spatial factor in small groups. *Journal of Abnormal and Social Psychology*, 54:269-72, 1957.
- A. Kendon. Some functions of gaze-direction in social interaction. *Acta Psychologica*, 26:262-63, 1967.
- M.L. Knapp and J.A. Daly. *Handbook of interpersonal communication*. SAGE Publications, Thousand Oaks, CA, 2002.
- Maxis, *The Sims*, Electronic Arts 2000.
- D. McNeill. *Hand and mind: what gestures reveal about thought*. University of Chicago Press, 1992.
- D. McNeill. *Gesture and thought*. University of Chicago Press, Chicago, IL, 2005.
- McNeill Lab. *Center for Gesture and Speech Research*. <http://mcneilllab.uchicago.edu/> (from February 2006).
- M. Masahiro. *Bukimi no tani [The Uncanny Valley]*, *Energy*(7):33-35, 1970.
- P. Olivier. Gesture Synthesis in a Real-World ECA, *Lecture Notes in Computer Science*, 3069:319-322, Springer, 2004.
- B. Reeves and C.I. Nass. *The media equation: how people treat computers, television, and new media like real people and places*. Cambridge University Press, New York, 1996.
- Rockstar Games. *Grand Theft Auto 3*. Edinburgh, Scotland, 2002
- Valve Corporation. *Half-Life 2*. Valve Corporation. Bellevue, Washington, 2004
- J. Weisz and G. Adam. Hemispheric preference and lateral eye movements evoked by bilateral visual stimuli. *Neuropsychologia*, 31(12):1299-1306, 1993.
- M.P. Wilbur and J. Roberts-Wilbur. Lateral eye-movement responses to visual stimuli. *Perceptual and motor skills*, 61:167-177, 1985.