Synthetic Character Fidelity through Non-Verbal Behaviour in Computer Games

Philip Heslop, John Shearer, Patrick Olivier

Informatics Research Institute
University of Newcastle upon Tyne
Newcastle upon Tyne
NE1 7RU, United Kingdom
philip.heslop@ncl.ac.uk, john.shearer@ncl.ac.uk, 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. Artificial Intelligence in games has historically focused on providing a challenging opponent for a player and narrative development. Scope exists to increase the fidelity of synthetic characters throughout the game to create a more immersive game play experience. This requires both visual and behavioural fidelity, and while graphics are nearing photorealism, synthetic characters’ behaviour is still unrealistic. Non-verbal behaviour of synthetic characters has to date received little attention and so the scope and participants of non-verbal behaviour requires identification. We review the range of spatial and task scenarios relevant in a game context, then identify categories of non-verbal behaviour and go on to summarise their role in communication and propose their incorporation in the design of non-player characters. Finally we review how non-verbal behaviour of synthetic agents might increase immersion for a player and identify interaction techniques that might facilitate non-verbal communication with players and non-player characters alike.

Introduction

The goal of the modern game designer is to produce an interactive world that engages a player to the extent that they feel like they are part of the environment rather than an external force that merely provides influence to how events in the game world transpire. This level of involvement is not merely a function of ever more realistic graphics or super intelligent enemies. Immersion is accomplished by designing the game world as a whole, whether the setting is realistic or abstract. One of the areas games have yet to address fully is interaction between the user and synthetic characters, and between the synthetic characters themselves. In many situations, characters in computer games do not behave like we would expect real people to behave. In order to
produce a more immersive experience for the player, characters should act in a manner that appears real. Interaction between characters in modern game worlds takes a variety of forms, some of which are already richly explored. Modern games have challenging enemies that contest with the player, have believable collaborative interactions when the action demands, and when narrative requires it, as a conversational tool to progress the story. Games are less sophisticated in how they deal with the indigenous population of non-player characters who populate game worlds without any direct involvement in action sequences or narrative. High level coordination of non-player character behaviours (typically through the characterisation of state space) has been the focus and less 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. When called upon to interact in this way, characters often exhibit unrealistic behaviour or simply omit realistic behaviour. These low level behaviours are actually a communication task: A character who exhibits them is trying to communicate some aspect of their internal state – in the case of artificial characters some element of their beliefs, desires or intentions. These encounters between characters and the player (and indeed observed interactions between characters) can be categorised by task context and are affected by range. In identifying these categories and exploring the effect of distance exhibited by real humans, we can ascertain the requirements for synthesizing these behaviours in a game character. Once these aspects of character behaviour have been identified we can further examine the shortfall in current games by looking at specific examples of how character interaction is currently implemented. This then allows us to suggest some possible improvements to the implementation techniques used in characters that would incorporate the suggested improvements.

Finally we offer some suggestions for facilitating the interaction between the player and game characters that allows the player to utilise this richer set of interactions when encountering other characters in the game world.

Task Context & Proxemics

As already described, communication provides information as to the beliefs, desires, and intentions of a character or some other aspect of the internal state of a character - showing a characters cognitive, emotional or physical condition. Characters who occupy the same shared space can exhibit verbal and non verbal communication, which can be categorised into four different task contexts, (Knapp and Daly, 2002, p177):

- Cooperation – Communication occurs so that characters can perform a task together, or work towards a common goal.
- Coaction – Characters exist in the same vicinity, and are aware of each others presence.
- Competition – Characters perform tasks at the expense of each other, or work towards conflicting goals.
- Conversation – Characters communicate in order to entertain or pass on information.
The method of communication varies across these contexts (and also with the physical proximity). When an encounter between characters, be they player or non-player, is initiated, the communicational context is established.

There are several ways context can be established without explicit communication. If the intention is strong then the context can be implied by the actions of the character without any explicit communication attempt. Context can also be implied by appearance, which although not a direct communication it can be seen as a premeditated attempt to convey some information. The motion of characters can also indicate the intended context of communication. Characters moving together look like they are cooperating, or at least in a group, while characters moving towards each other can look like they are in competition. A character that is motionless can still imply a context, especially if motion would help or hinder another character. For example a character not moving out of the way in a narrow corridor when another character wishes to pass could imply competition.

Explicit communication provides a more obvious cue to intention. Verbally a character can simply tell another their intention, or the intention can be implied by the manner in which the character speaks. This aspect of communications is well established in computer games, though it is still often lacking in incidental encounters that are not part of the narrative.

Non-verbal communication is not as well utilised in computer games. Facial expression, eye movement, body movement and gesture are all used in normal communication but are often omitted or under represented in synthetic characters.

Use of the eyes is an important component of human-human communication. Direction of gaze is utilised in communication, 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:

- Regulating the flow of communication
- Monitoring feedback
- Reflecting cognitive activity
- Expressing emotions
- Communicating the nature of an interpersonal relationship [added by Knapp].

Synthesising these in a non-player character requires more 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, acting as a cue for turn-taking and highlighting grammatical breaks, conceptual unit breaks, and the ends of utterances (a sequence of speech separated from another by a marked gap), while length of gaze shows a desire to change the level of interaction. These glances also allow feedback on the interaction by monitoring the reactions of the other person, although a synthetic character would not literally do this, it expected during conversation and so should be part of the synthesis.

Gaze can also be used to convey some elements of the internal state of a character. Cognitive load (trying to process difficult or complex ideas) can lead both listeners and speakers to look away, the averted gaze reflecting a shift in attention from the external to the internal. There is evidence that the eye gaze direction under this condition changes with different forms of cognitive load, linked to the active hemisphere of the brain (Ehrlichman and Weinberger, 1978; Weisz and Adam, 1993; Wilbur and Roberts-Wilbur, 1985).
It is not simply what a person is looking at, but people use the eyes and facial expression to communicate a wide variety of emotional states, such as downward glances indicating modesty, wide eyes indicating frankness (or wonder, naivety or terror depending on other visual cues) (Knapp and Daly). Basic emotions such as surprise, fear, disgust, anger, happiness or sadness can be expressed through the eyes. 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 can indicate the status (or the perceived status) of the communicants (Heam, 1957; Efran, 1968) or it can indicate the level of intimacy between the communicants. Lovers, or mothers with babies, have extended periods of mutual gazing.

Gesture is another communication form that becomes important at the personal and intimate ranges. Movement that serves to communicate, mainly involving movement of the hands and the head, can be classified as gesture:

- **Emblematic** - (gestures with specific meanings that occur without speech) - are well defined in both their form and meaning and are therefore readily synthesised with current animation techniques. 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).

- **Spontaneous** - (hand and head movements that occur with speech) - performed in synchrony with the speech and is generally made with the head or hands, or if that is not possible then any available body part (or even the whole body) for example pointing with a foot when one’s hands are full. Recent research on spontaneous gesture indicates a link to other cognitive processes (McNeill Lab, University of Chicago, 2006; McNeill, 2005).

Spontaneous Gesture is tied closely to (but is less constrained than) the underlying speech and both speech and gesture are widely believed to be generated from a single underlying conception (McNeill, 1992). Spontaneous gesture can be complementary, supplementary, or contrastive to the speech, it can re-iterate or emphasise the speech, add information to the speech, or communicate something contradictory (or slightly different) from the associated speech. McNeill identifies five categories of spontaneous gesture:

- **Emphatic** - (also known as beat gesture, or baton gesture) provides emphasis. It is used to highlight parts of speech, key phrases, words or even phonemes, and consists of two complementary movements (up/down, in/out, left/right), with the transition being the point of emphasis.

- **Deictic** - pointing actions that refer to an object or objects. They can reference more abstract concepts, 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).
- Cohesive - connects related parts of conversation that are temporally separated. For example, when listing items people often provide an emphatic gesture on each item, connecting them through repetition of the same gesture.
- Iconic – pictorial representation of a concrete concept, such as an object or acting out a specific action that is being described (McNeill, 1992).
- Metaphoric – pictorial representation of an abstract concept as a physical object, such as assigning an object to ‘a great idea’.

In order for a character to represent objects though gesture, they would need some concept of the object’s shape or movement, hence iconic gestures are probably the most complex gestures to synthesise. The gesture can also change depending on 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 the size of the cup, the gesture would be distinctly different.

Gesture is also used much as eye gaze to regulate communication flow (to indicate turn-taking for example) (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 and frequently coincides with flow regulation signals in other channels.

Finally, there are the general movements of a character, which while having no communication purpose can still be observed and interpreted by other characters. These movements, known as self-adaptors, include motions such as stretching or scratching 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 sometimes interpreted as a flirtatious behaviour. Characters can use also their posture or body position, as well as props, to give out specific messages. For example a character holding their arm out in front of them could be communicating ‘do not pass this way’, and this can be reinforced if the character is holding a weapon, while a character with their arms open can be interpreted as a friendly action, inviting closer contact. Though not gestures in the way described above, these body movements still convey information and are also a method of communication.

Character communication is also affected by the proximity of the communicating parties. This closeness and arrangement of the self in physical world compared to others is known as proxemics (Hall, 1966). Each character in the environment has its own personal space and territory. It is difficult in some genres of game for the user to appreciate their own personal space and territory, as there is no visual representation of the user. Increasing the immersive experience should improve this, as the reactions and behaviour of other characters will give the appropriate cues. The impact of proxemics will increase as characters become more sophisticated and technology allows more realistic representations.

Proxemic behaviour also exhibits some of the largest (but consistent) cultural variations, especially in conversation. Concentrating on UK/US culture (Hall, 1966) proxemic spaces are categorised as:
- 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, verbal communication at this range is almost non existent.

These proximal differences stem from the idea of personal territory. Convincing non-player characters therefore would use proxemics as people do – to be socially correct, a character would have different levels of proximity that they might use depending on who they were interacting with, and how well the character knows them, in line with the above categories. Violation of these rules is socially incorrect and in itself implies intention of a character. For example, inappropriate invasion of personal space can be intimidating or flirtatious. (Of course characters from different cultural backgrounds might inadvertently break social conventions). Certain conditions can override social convention however, such as a task that requires physical proximity such as tending to a wound. Even in these cases however a character would be given tacit permission to approach by the wounded character’s cries of pain or a direct plea.

The idea of territory can be extended if there is the notion of ownership of certain parts of the environment. The environment can be parameterised into primary, secondary and 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 ownership of a territory can be temporary and can lead to different behaviour towards that object or space while a character ‘owns’ it (Knapp and Daly, 2002).

**Case Studies**

Within the many genres of games currently available, improvements in character behaviour will have the greatest impact in games where the player has direct encounters with AI controlled characters – games which have some representation of the player either in first person or as an in game avatar. Therefore we limit our investigation to games of this type.

One of the most influential first person games of recent times has been Quake 3 Arena (ID Games 2001). The game is based around a tournament where the aim is to frag or kill opponent characters, which may be other human players or AI controlled bots. When playing against computer controlled characters, they show their intent to compete with the player and each other directly through their actions - that is they attempt to frag their opponents. There is also some generated in game messages that the characters output to the chat window in a very basic attempt to engage the player. However these messages have no effect on other characters, nor do the simple gesture commands that play animations when certain keys are pressed. In fact the game has no observable character communication, even when AI characters are working in a team (such as in capture the lag games). Nor is there communication with the player other than direct action.

Grand Theft Auto 3 (Rockstar Games 2002) is another influential game. It uses a third person, ‘over the shoulder’ perspective. The game is partially plot driven, with missions that the player can complete introduced by cut-scenes. During the game itself the player has a great deal of freedom to wander around the city, which is heavily
populated with pedestrians and gang members. The populace is able to react to the actions of the player not only with aggression as in Quake 3 but also with fear, as civilians will run away if the player fires a weapon. Although a step above the one dimensional characters in Quake 3, the fidelity of the characters is still lacking. Characters rarely acknowledge each others existence, except to avoid collisions or to attack a rival gang member. They do not attempt communication with the player unless they are threatened, and other characters will go about their business while the player attacks someone in front of them.

Half Life 2 (Valve Corporation 2004) is a more subtle game that is more story-driven and so has many more instances of character communication. 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 will move around to maintain eye contact during the conversation; respond to the players actions; and show expressions and gestures appropriate to the game-play, thus attempting to engage the player with the narrative. Aside from narrative events, however, character communication is still limited. While the characters within Half-Life 2 appear to have desires, personality, and intentions, their actual behaviour is 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. These complex behaviours attempt to fulfill the role of the traditional cut-scene, but without the break in game-play. However, the integrity of the non-verbal behaviour relies on the skill of the animators in coordinating the non-player character’s action with the game-play and the character’s speech. Outside of narrative events, communication is limited. For example when a non-player character and the player collide, 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 their ability to break them) both undermines the engagement of players with the game and limits they expressivity. For example, walking into another character and the social conflict that arises may serve to further or undermine a player’s ends.

Co-Action, Competition, Cooperation and Conversation in Games

There are many possible improvements to non-player characters’ behaviour utilising what has been observed in human behaviour with regard to task context and proxemics that would lead to a more realistic environment.

Co-Action in games is where the engagement of players would be most improved by incorporating non verbal behaviours in non-player characters, as it occurs often and at times in the game play when the player is less distracted by action or narrative content – therefore having more attention to detail - It is also the area which most current games neglect when implementing their synthetic characters. When characters have to negotiate the same space they have the most opportunity for incidental communication. For example collision avoidance would be much improved if non-player charac-
ters used gaze and gestures to indicate their intended path, or show which side they intend to pass. In current games characters walk around in a robotic like manner and do not seem alive. To an observing player, characters that incorporate these non-verbal behaviours into their actions produce a more believable population of realistic characters in the environment.

Characters in games clearly signify when they intend to compete through direct action against those they compete against, but fail to incorporate subtler behaviours. If a character could communicate their emotional state through non verbal behaviour, an observer could tell if a character is getting angrier and is poised to attack, or if a character is edging towards a common goal (by noting gaze direction and gestures). This in turn would give much more choice to a player, maybe it is wiser to placate an angry character and not antagonise them, or a competing character’s behaviour might indicate the location of some hidden common goal. Giving the player the extra cues that interpreting another character’s behaviour would bring would add to the immersion in the game experience.

Similarly with co-operation, non-player characters exhibit relatively good communication when acting directly, but are less effective at other times. Utilizing gestures and gaze direction, characters could communicate information that would help achieving a common goal. In its most basic sense it could be simply pointing, but could also include gestures to aid the description of objects or people in the environment. Observing characters cooperating or competing with each other in this way shows their relationship and allegiances. It could be an important tool in establishing characters personalities to the player.

Conversation rarely occurs outside of narrative events, and usually only a specific message is communicated. Outside of these events, characters rarely even speak to a player or each other, and they don’t use many gestures or other non-verbal communication. Including conversational gestures and expressions would not only increase the fidelity of narrative events, it would make incidental conversations more believable.

This is also true of observed conversations between non-player characters. The emotional state of characters can be portrayed, revealing relationships and hierarchies as well as simply making the characters look more real.

These relationships and hierarchies are important in all interactions. Characters should react differently depending on familiarity, social standing and impression they have of another character. A character will greet a friend differently from an unknown authority figure for example. The different way characters react is a big cue as to their ‘personality’. A character can be categorised as friendly if they greet a certain set other characters warmly, or simply gregarious if they greet anyone warmly. The appearance of social behaviour is imperative for producing a realistic population of characters – therefore vital in producing a highly realistic, immersive environment.

**Synthetic Character Implementation**

Synthesising these non verbal behaviours have not been attempted previously in games is not only due to technological limitations. Characters have not been implemented with the knowledge of how and when these behaviours should occur. Characters simply do not have enough information in their internal state to act on, only hav-
ing parameters that directly affect their actions and not the manner in which these 
actions take place.
A much richer set of beliefs is also required with the belief desire intention model of 
the character. Beliefs should deal with how actions should take place, not just when or 
why. For example a character who believes they strongly like another character 
would have a set of beliefs that coincide with how to behave towards a another char-
acter given that premise, how close to stand when conversing, how eye movement and 
gaze can illustrate their affection, which gestures to use when speaking etc.
The character also requires feedback on their actions, perhaps the other character does 
ot feel the same way and uses their own set of beliefs to back off slightly or be more 
aloof. As a character cannot ‘see’ in the sense that a player can observe the other 
characters in the environment, a mechanism for transmitting information about how a 
character looks needs to be implemented. Characters need to know how another char-
acter is acting as much as they need to know what they are actually doing.
There is a technical issue still to address however, about what appears to be the large 
number of animations that would need to be produced to show all the gestures and 
movements that would be needed. However many of these could be produced pro-
grammatically, with a small set of animations linked together to form a certain ges-
ture, and using variable speeds and degrees of motion to add an emotional element. 
For example a sense of urgency can be created by increasing the speed of a gesture, or 
anger could be simulated by increasing the speed and the range of movement. It can 
already be seen from the example of Half Life 2 above that gaze direction can be al-
tered dynamically when conversing, this head movement could be further utilised to 
synthesise the other uses of gaze direction outlines previously. The same could be 
done with arm or leg movements so that gestures appear spontaneous or at least di-
rected appropriately. More general gesture generation, especially for iconic gesture is 
an open research problem in psychology (when people do it) and computer science 
(how to simulate it).

User Interaction Implementation

Whilst we can imagine a situation whereby non-player characters have sophisticated 
cognitive models and the ability to both synthesise and interpret non-verbal behav-
ior, there is no mechanism for the player to communicate non-verbally with either 
the non-player characters or other players. Perhaps with the advent of simple motion 
tracking systems being part of the standard interface with games, such as Nintendo’s 
Wii controller (that operates somewhat like a three dimensional pointing device) 
(Nintendo 2006), this barrier in communication can be overcome.
Present-day technology does exist that 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 dil-
ation, facial expression, vocal behaviour, voice, and a variety of other biometrics - 
although some elements are extremely transitory and require a very high sampling 
rate to capture. Unfortunately, such data can only be collected accurately using spe-
cialised (expensive) invasive equipment, but it is yet to be determined if this level of 
accuracy is required for interacting with non-player characters, and perhaps game
controllers like Nintendo’s Wii controller will be sufficient to allow a wide range of novel interactions. There are also under-explored interface technologies that are available now that could aid in this communication task, such as 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:

- 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.
- 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. Gestures are an automatic result of the ‘mood’ of the player’s character. For example a certain keyboard command would generate a specific gesture.
- 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.

**Conclusion**

Synthetic characters in computer games play an important role in providing the challenge and narrative tools to enhance the experience of game play. However, they are currently underused as a tool for providing a believable and immersive experience outside of these specific areas. In order to create characters with behavioural fidelity we need to consider the aspects outlined in this paper. Human-like characters need to behave human-like at all times, not just when called upon to provide narrative or to be a challenging opponent. In order to behave more human like, characters need to express themselves to the player and other characters in a believable manner, including the use of non-verbal communication techniques we all recognise from our everyday interactions with each other and have more complex models of other characters and their relationships. Some of these communication paradigms are complex and are only now being understood on a psychological level, and may be difficult to synthesise, but when these obstacles are overcome, the reward will be a much more immersive and fulfilling game experience.
References


Rockstar Games. *Grand Theft Auto 3*. Edinburgh, Scotland, 2002

