Autonomous Secondary Gaze Behaviours

Marco Gillies[†]; Neil Dodgson[¥]; Daniel Ballin [‡]

[†] Department of Computer Science, UCL@Adastral Park, Ross 1, Adastral Park, Ipswich, IP5 3RE, Suffolk, U.K.

[¥]University of Cambridge Computer Laboratory, William Gates Building, JJ Thomson Avenue, Cambridge CB3 0FD, UK

[‡] Radical Multimedia Laboratory, BTexact Technologies, Ross 1, Adastral Park, Ipswich, IP5 3RE, Suffolk, U.K.

[†]marco.gillies@bt.com; [¥]neil.dodgson@cl.cam.ac.uk; [‡]daniel.ballin@bt.com

Abstract

In this paper we describe secondary behaviour, this is behaviour that is generated autonomously for an avatar. The user will control various aspects of the avatars behaviour but a truly expressive avatar must produce more complex behaviour than a user could specify in real time. Secondary behaviour provides some of this expressive behaviour autonomously. However, though it is produced autonomously it must produce behaviour that is appropriate to the actions that the user is controlling (the primary behaviour) and it must produce behaviour that corresponds to what the user wants. We describe an architecture which achieves these to aims by tagging the primary behaviour with messages to be sent to the secondary behaviour and by allowing the user to design various aspects of the secondary behaviour before starting to use the avatar. We have implemented this general architecture in a system which adds gaze behaviour to user designed actions.

1 Introduction

In general when a user controls an avatar they will give a sequence of commands that will be transformed into the avatar's motion. The method of giving commands will vary greatly, from simple mouse and keyboard interfaces to various forms of body tracking. Though purely text based interfaces to avatars are rare, there is one exception, which is that some avatars are controlled by the user typing text into an interface, that the avatar then speaks. The level of control also varies greatly, the system might map the users motion directly onto the avatar or the input might be a much more discrete set of higher level commands, for example a command to walk to a particular position.

What these methods have in common is they are incomplete. They cannot specify the entire behaviour of the avatar. Even full body tracking systems cannot capture the full detail of the user's motion and expression (of course this would normally not be desirable, part of the appeal of using avatars is that they add some thing new to the action of the user whether it is in graphical appearance or behaviour). As discussed above this leaves various aspects of the avatars behaviour that must be determined by the system.

2 Primary and Secondary Behaviour

We divide the avatars behaviour into two types: primary behaviour that is explicitly specified by the user input, and secondary behaviour that is automatically generated by the system. Primary behaviour should correspond to large scale goal directed actions, such as moving around or manipulating objects. They should be the sort of actions that people consciously decide to do, thus making them easy for users to specify. Secondary behaviour on the other hand should be smaller scale, and should tend to correpsond to more sub-conscious behaviour such as non-verbal communication (or at least behaviour that is not the main focus of the user's action). For example, a primary behaviour would be invoked if the user requests the avatar to pick up a telephone and to start talking. Secondary behaviour accompanying this might be a head scratch or fiddling with the telephone cord.

We envisage that secondary behaviour should be controlled by a number of independent behavioural agents. They would produce their behaviour continually, in parallel with any user controlled behaviour. Though the secondary behaviour is not explicitly specified by the user it should be influenced by the user's input. In particular the details of the secondary behaviour should depend on what primary actions the avatar is performing.

2.1 Motivation

However graphically appealing avatars can be they will not become successful if their animation and behaviour is not compelling. It is thus very important that the behaviour of avatars is expressive. For this to be possible the behaviour must capture the nuance and complexity of human non-verbal behaviour. This is not possible solely through the user controlling the avatar directly. Most input devices will not provide enough information to animate details such as the character's gaze behaviour, gesture or facial movements. A mouse and keyboard interface is far to impoverished to control all of these in real time. Even a full body tracking system can miss significant information. Also full body tracking is not suitable for cases where the user does not want their body language directly mapped onto the avatar, for example, to hide their feelings or to make the avatar's behaviour more stylised or more expressive. Directly controlling the character's expressive behaviour would also be a large cognitive load on the user who is likely to want to concentrate on the task in hand. They would probably not pay much attention to the details thus resulting in avatars that are not expressive in practice. Finally, even if the user could control the avatar's expressive behaviour directly much of the behaviour is sub-conscious and so the user might not know how to produce appropriate animations if they are not a skilled animator. All of these reasons indicate that expressive behaviour should be generated autonomously. However, it is still important that the behaviour should be relevant to what the character is doing, i.e. what the user has commanded it to do. Thus we divide the avatars behaviour into primary and secondary behaviour and ensure that the secondary behaviour is autonomous but influenced by the primary behaviour.

2.2 User design of secondary behaviour

One aspect of secondary behaviour that we consider very important is that the user should be able to control many aspects of the avatar's secondary behaviour. This should be done by giving the user tools with which to shape the secondary behaviour before starting to use the character. This allows us to harness human creativity in the process of producing expressive behaviour. People can be excellent at creating the subtleties of human expression and the ability to harness this can add a lot to an expressive agent. This form of user input can also provide a large degree of individuality to an avatar. It can be particularly important when the avatar or environment is stylised so that what is needed is not realistic behaviour but some form of stylised behaviour (cartoonish, exaggerated etc.). As this design is done before the avatar is used it does not force the user to spend time controlling the secondary behaviour in real time but still allows them control over the avatars behaviour.

3 Previous Work

Vihljámsson and Cassel (1998) discuss the importance of autonomous behaviours for avatars, which are equivalent to what we call secondary behaviour. They use these autonomous behaviours for conversation in their BodyChat system. Similar ideas are present in BEAT (Cassell, Vilhjámsson and Bickmore 2001), which generates conversational gesture from text. This work differs from ours in that we are looking at graphical control of physical, non-conversational primary behaviour. Another related system is the multi-level control of Blumberg and Galyean (1995), this allows the user to control some aspects of a characters behaviour, at various levels of abstraction with the system generating others, for example, wagging the tail of their dog character. Finally, somewhat related is the current interest in transforming pieces of motion to express new emotions or personality (e.g. Rose, Bodenheimer and Cohen 1998, Amaya, Bruderlin and Calvert 1996 and Polchroniadis 2000). Though these do not really count as autonomous behaviours they do add secondary features to a motion.

4 General Architecture

Figure 1 gives an overview of the architecture that is being proposed for primary and secondary behaviour. The primary behaviour is controlled by direct user commands. The Secondary behaviour is a separate module (or set of modules) that is not directly influenced by user input and which acts to a large degree autonomously.

The secondary behaviour module is controlled by a number of tags attached to the primary behaviour. These tags are attached to particular points in the primary behaviour and indicate that a message should be sent to the secondary behaviour when that particular point occurs in the primary behaviour. For example, in a conversational system a tag could be attached to the point at which the avatar stops speaking and this could result in various secondary actions being requested from the secondary behaviour module, for example, looking at the conversational partner.

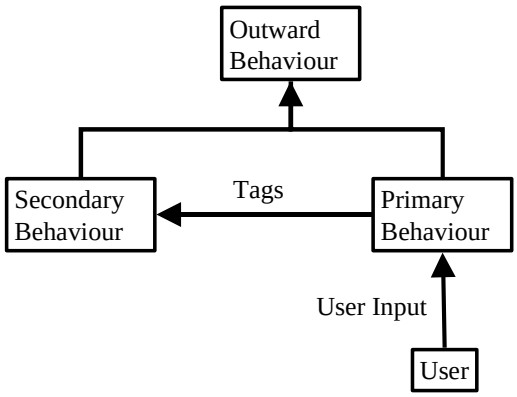

Figure 1: The relationship between primary and secondary behaviour.

These tags can be controlled by the user designing the secondary behaviour. There are two aspects that can be controlled, the points in the primary behaviour at which the tags are placed and the message that is sent by a tag. There are also two types of user who will edit these tags. The first will be designing the secondary behaviour system in general. They will design how the secondary behaviour relates to the primary behaviour and so will mostly add tags to the primary behaviour. They will also edit the messages contained in the tag to some degree. The second type of user will create the secondary behaviour for a particular avatar and will add details to the system designed by the first user. They will mostly edit the messages contained in the tags. For example, the first user might add a tag requesting that the avatar should look at the partner at the end of an utterance while the second user might indicate whether this should be a brief glance with just the avatars eyes or whether the avatar should orient itself towards the partner with its head and shoulders and look at the partner for a longer time.

5 Example: Eye Gaze

We have implemented an example of this general architecture for generating eye gaze while an avatar obeys commands given by the user. The simulation of gaze behaviour has been studied extensively for conversation, for example, Vilhájmsson and Cassell 1998, Colburn, Cohen and Drucker 2000 and Garau, Slater and Sasse 2001. We have looked instead at gaze behaviour in non-social situations. This has been studied by Chopra-Khullar and Badler but they did not investigate in detail how to integrate simulation of gaze with user control of the avatars actions. We focus on creating tool by which a user without programming knowledge can create both primary actions that the avatar can perform as requested by the user and gaze behaviour that will accompany these primary actions.

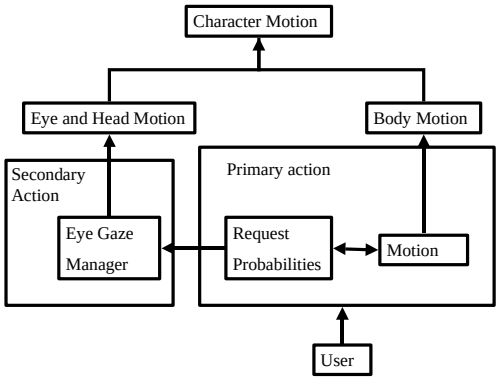

Figure 1: The Gaze Behaviour

5.1 Primary behaviour

The aim of our primary behaviour is to supply a framework for actions that can both be easily designed by a user and also easily invoked by the user in real time. When creating an action the user starts with a piece of motion, this motion is a specific example of the action to be designed. The user then specifies targets of the motion (objects that the character interacts with during the motion). For example, a target for a drinking motion would be a cup. Targets can also be objects that the character looks at but does not touch, for example, the character might be drinking in a cafe talking to a friend. This friend might be a target so the character would look at her.
The user can invoke these actions by specifying targets, which is just done by clicking on them. When an action is invoked the original motion is transformed so the character interacts with the targets correctly. For example, if the action consists of picking up and drinking a cup the motion will be transformed so that while picking up the cup the hand will move to the correct position of the specified cup rather than the position that it moved to in the original motion. The motion is transformed with techniques similar to those used by Polichroniadis (Polichroniadis 2000). The full details of the method is described in Gillies 2001.

5.2 Gaze behaviour

The secondary behaviour consists of gaze shifts which are controlled by an eye gaze manager (described in more detail in Gillies and Dodgson 2002a). This receives requests for a gaze in a particular direction or at a particular object from the primary action. These requests can be of

different types, for example, an immediate request causes the avatar to look at the target as soon as the manager receives the request while a monitor request just causes the avatar to look at the target occasionally until it is told to stop by another request. The eye gaze manager animates the avatar as looking at the target of these requests. When no requests are sent the manager continues to produce gaze behaviour, generating requests itself.

The eye gaze manager is controlled by a number of parameters that influence the avatar's gaze behaviour. For example, observing people we noticed that they vary their horizontal angle of gaze but kept their vertical angle relatively constant. Thus we introduce two parameters to control the characters behaviour, a preferred vertical gaze angle and a probability of maintaining this angle. These parameters can be altered by the user to change the avatars behaviour. These parameters can be set in advance and allow the user to alter the avatars behaviour in general without having to do anything while actually using the avatar.

Some parameters can also be set for individual requests. For example, the length of gaze is controlled by a number of parameters for the character, however, it can also be set by parameters of a request. This allows two levels of control. The parameters of the gaze manager can change the avatar's gaze behaviour in general, while the request parameters change how it behaves in particular circumstances. These parameters allows different avatars to have a range of different gaze behaviours using the same architecture.

Finally, the avatar's gaze behaviour can be influenced by external events, for example, various objects can capture the avatar's attention, for example, moving objects. Also when there are no requests the avatar will look at objects in the environment tagged as interesting.

5.3 Tagging primary behaviour

As described in section 4 above the primary behaviour is tagged with messages that are sent to the secondary behaviour module at various points in the primary behaviour. These messages are to ensure that the secondary behaviour produced is appropriate to the primary behaviour. In this case the messages consist of eye gaze requests.

When the user creates a primary action from a piece of motion she divides the motion into a number of periods. These periods correspond to meaningful sub-sections of the motion. For example, for a motion of picking up and drinking a cup of coffee the periods might be, reaching out towards the cup, picking up the cup, bringing the cup to the avatar's mouth, drinking and putting the cup back down. Tags can then be placed at the start of these periods.

The user sets the tags by specifying that particular types of requests can be sent out at the start of each period. These requests will be to look at one of the targets of the action. Various parameters of the request can also be specified, for example whether it should be a short glance or a longer gaze; whether the avatar should move its head, and how often the avatar should look at a target if the request is a monitor request. If the gaze behaviour produced were identical every time the action was invoked it would seem very mechanical if the action is invoked a number of times. To prevent this the user actually specifies probabilities of sending a particular type of request. When an action is invoked requests are generated at random using these probabilities. The probabilities are set using a number of sliders as shown in figure 2.

When the user first creates the action they can set these probabilities, however, it would also be desirable to be able to alter them to create different personalities for different avatars. Users can therefore alter the probabilities later. We separate these two editing stages by making the original designer's edits effect what later edits are possible. These original edits might just be default values that users can later change. However, some values might be inappropriate, for example, when drinking it would always be appropriate to keep the avatars head still so the original designer might want that edit to be permanent, so that users cannot later change it. The original designer can indicate that an edit is a fixed value and so it will not appear in the users dialog-box for later editing. The original designer can also set particular edits as minimum or maximum values.

5.4 Results

Figures 3 and 4 give examples of actions with eye gaze attached. The first is of an avatar drinking from a can. The underlying gaze parameters are set so that the avatar has a tendency not to look around itself and to mostly look downwards when there are no explicit requests. There are two requests tagged to the actions. The avatar looks at the can before picking up and then at the other avatar at the last frame, this time just glancing and moving its eyes without turning its head. The behaviour of not looking at the other avatar in general and when looking doing so without a head move might indicate avoiding the gaze of the other avatar. The Second example is of an action where the avatar picks up an object and puts it down somewhere else. Here the avatar looks around itself more. There are two tagged gaze requests, to look at the object as it is picked up and at the shelf as it is put down. This time, when the character does not have a request

between the other two it looks at a location in the distance.

6 Further Work

Both secondary behaviour in general and our particular system have a large potential for further development. We have only implemented one particular type of secondary behaviour applied to one particular type of primary behaviour. Thus there is much potential for exploring new types of secondary behaviour, for example, gesture or facial expression. Also there is a lot of potential for applying these to other types of primary behaviour, for examples, conversation or other behaviour patterns that are too complex for to fit into the framework described here. The BEAT system aims at attaching gesture to speech (see Cassell, Vilhjámsson and Bickmore 2001). The authors have also investigated adding eye-gaze to a more complex primary behaviour, navigating an environment (described in Gillies 2001 and a forthcoming).

The tool we have described here is still a prototype and needs to be made more robust and tested by creating a wider range of actions and performing user tests. In particular we would like to develop it into a tool that can be used in shared virtual environment and assess people's perception of avatars using our secondary behaviour.

One aspect that we would like to improve is the user interface for adjusting the various parameters of the secondary behaviour. These allow the user a degree of control over how a particular avatar performs its gaze behaviour. However, these are currently edited using a large set of sliders that directly effect the parameters, some of which are rather unintuitive, we would like to provide a more sophisticated and intuitive design tool.

Though this model of eye gaze is reasonably general it is not quite sufficient to model the nuances of interpersonal eye gaze in social situations, we would therefore like to include more heuristics for social situations.

7 Conclusion

We have explored the idea of secondary behaviour as a semi-autonomous system that controls aspects of an avatars behaviour, leaving the user free to control more important aspects. We have described an application of these ideas to generating eye gaze. We think this has provided a good demonstration of our general architecture and are pleased with our initial results, however, we are keen to develop these ideas further.

Acknowledgements

The work descirbed in section 5 was done at Cambridge University Computer Laboratory and funded but the UK Engineering and Physical Sciences Research Council. The Continuation of this work is being funded by BTExact technologies. The authors would like to thank the members of the Cambridge University Computer Lab Rainbow research group, the radical multi-media lab and the UCL computer graphics and VR group, and in particular Tony Polichroniadis, for a conversation that sparked some of these ideas.

References

Kenji Amaya, Armin Bruderlin, Tom Calvert, Emotion from Motion, in *Proceedings of Graphics Interface '96*. pp. 222-229, 1996.

Bruce Blumberg and Tinsley Galyean, Multi-Level Direction of Autnomous Creatures for Real-Time Virtual Environments, In *Proceedings of ACM SIGGRAPH* 1995

Justine Cassell, Hannes Högni Vihljámsson and Timothy Bickmore, BEAT: the Behavior Expression Animation Toolkit, *in Proceedings of ACM SIGGRAPH* 2001

R. Alex Colburn, Michael F. Cohen and Steven M. Drucker, The Role of Eye Gaze in Avatar Mediated Conversational Interfaces *Microsoft research Technical Report* MSR-TR-2000-81

Maia Garau, Mel Slater and Angela Sasse, The Impact of Eye Gaze on *Communication using Humanoid Avatars, in Proceedings of ACM SIGCHI* 2001

Marco Gillies, Practical Behavioural Animation Based on Vision and Attention, University of Cambridge Technical Report TR522 2001

Marco Gillies and Neil Dodgson, Eye Movements and Attention for Behavioural Animation, under review *Journal of Visualization and Computer Animation* 2002 (a)

Tony Polichroniadis, High Level Control of Virtual Actors, *PhD Thesis, Cambridge University Computer Laborotory*, 2000.

Charles Rose, Michael F. Cohen, Bobby Bodenheimer, Verbs and Adverbs: Multidimensional Motion Interpolation, *IEEE Computer Graphics & Applications. 18 (5)*, pp. 32-40, 1998.

Hannes Högni Vihljámsson and Justine Cassell, BodyChat: Autonomous Communicative Behaviors in Avatars, *In Proceedings of the 2nd Annual ACM International Conference on Autonomous Agents*, Minneapolis, 1998

Daniel Ballin, Matthew Lawson, Mary Lumkin, Jo Osborne, Personal Virtual Humans: Inhabiting The Talk Zone and Beyond, *BT Technology Journal*, Kluwer, Vol(20) No1, January 2002.

Daniel Ballin, Ruth Aylett, and Carlos Delgado , Towards the development of Life-Like Autonomous Characters for Interactive Media, *BCS Conference on Intelligent Agents for Mobile and Virtual Media, National Museum of Film and Photography*, April 2001.

Robert Trappl, Paolo Petta, Creating Personalities for Synthetic Actors, *LNAI* Springer, 1997

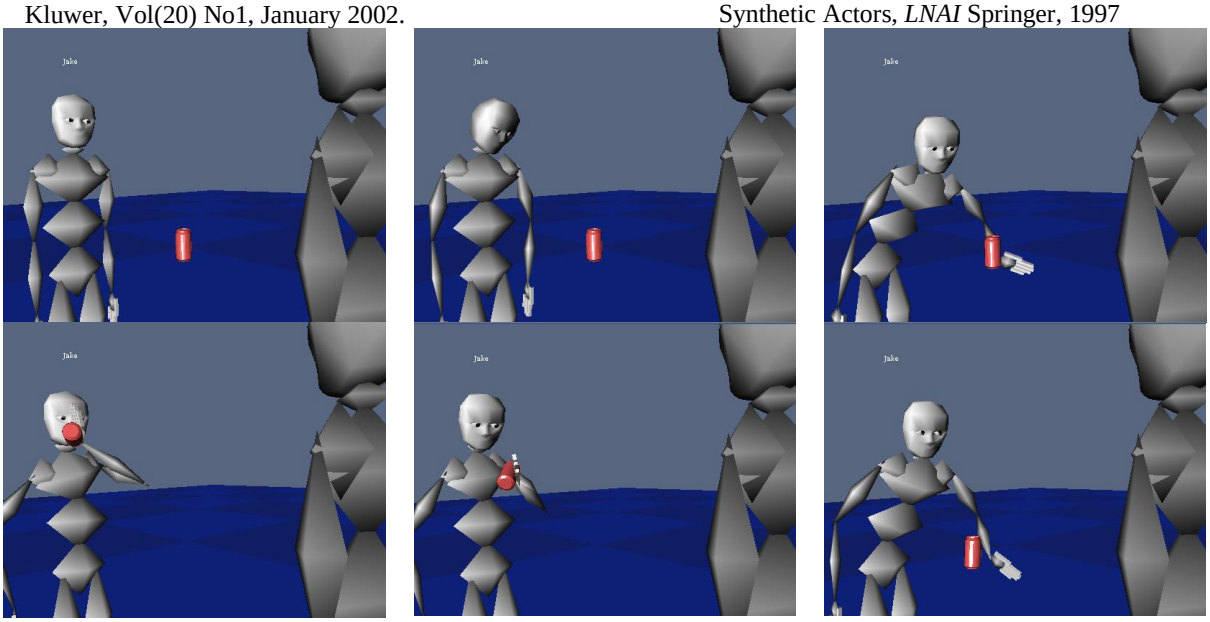

Figure 3: An action of an avatar drinking from a can

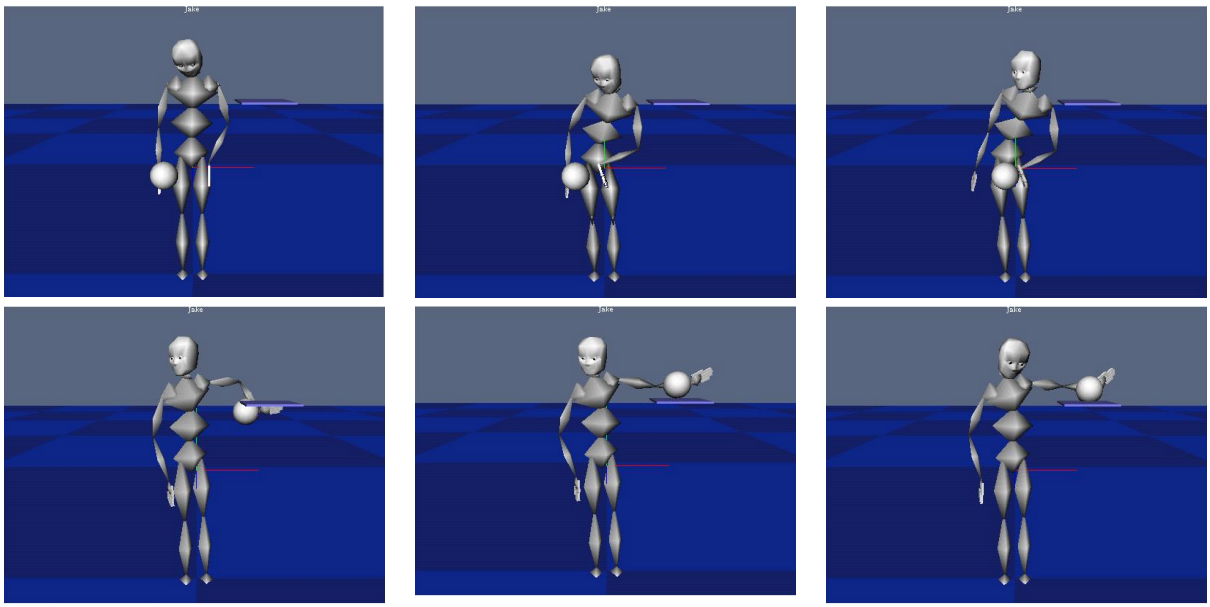

Figure 4: An action of an avatar picking up an object and putting it down somewhere else