Chapter 4

Social Interaction

People gained new knowledge by communication and social interaction. Virtual collaborative environments allow people communicate with others to exchange ideas, and interact with virtual objects and distant users in the same virtual environment. Social interaction, is when people communicate and interact with others in the real world, facial expressions and gestures stimulate this communication process. If the avatars only express its emotions through facial expressions, this creates an unnatural feeling for the users when they communicate and interact with other avatars in the virtual world. This is because gesture is another form of expressing emotions during communication and social interaction in the real environment, so attention must be given when modelling and animating the avatar body.

The first section in this chapter looks into the process of creating a simple full-body avatar suitable for a synthetic social environment and in how to animate the upper body movements of the avatar associated with facial emotions and other simple actions. Section 4.2 discusses the interaction process between different avatars and their virtual environment. Section 4.3 discusses briefly the effects of combining sounds in the virtual environment. Section 4.4 discusses the design of a simple state machine that will co-ordinate the interaction process between all the avatars in the synthetic social environment. The last section summarises this chapter.

4.1. Full body Avatar

When designing the body of the avatar, the body must be designed such that it is low polygon, able to perform animation, and the user must recognise that the model is a simplified human model. The reasons for these criteria are, when the number of avatars in the virtual environment increase, the computation and system resource usage will increase. By lowering the polygon structure of the body, the amount of system resource use will decrease, and more avatars can participate in the virtual environment. The body model must be able to animate body movements or simple gesture, to simulate real people interactions and communications in the virtual world. Although the body model is simplified, during interaction the users should feel or see that he/she is interacting with a human-like avatar and not a robot or a figure that cannot be identified easily.

4.1.1. Body Creation

The avatar's body is referred as a body model, because it is a representation of the human body. The body model is of geometric nature, because the model is a collection of components with well-defined geometry (the body consist of arms, hands, legs, feet, .etc) and interconnection between components exist (the hand is joined to the arm).
The advantage of representing the body as a geometric model is, because the inheritance and hierarchical geometric properties exist in geometric models, which similar to the human body e.g. when the arm moves, the hand move along with it.

The body model is decomposed into a collection of different parts, so that each part is modelled as an individual object e.g. the arm is separate into upper arm and lower arm. The reasons for modelling each limb as a separate object is, because the desired joint angle for each body part are determined more easily and the body with all the limbs can be structured hierarchically.

When drawing the human body artist usually first draw the head or the body and then the rest of body [100]. This is because the body must be position correctly in the picture first and then draw the rest of the body parts according to the position and size of the body. This is known as the top-down construction process.

In hierarchical modelling, the body model is created by a bottom-up construction process. This means that the lower-level body (e.g. Hand) parts are modelled first and served as building blocks from the higher-level body parts (e.g. Arm). The higher-level body parts are modelled after the lower-level body parts, if the body parts are out of proportion, they can be scaled to the correct size. Although the body model is constructed in a bottom-up manner, it does have one top-down characteristic in terms of colour, e.g. the hands has the same colour as the arms.

The body model can be simplified by modelling the clothing of the avatar as part of the body. The skeleton of the avatar is not modelled and complex limbs are further simplified by polygon reduction. E.g., the human hand is the most complicated limb on the human body, because it consists of many joints and folds. Instead of modelling each finger with joints, the fingers are modelled as a single polygon box (Figure 4.1).

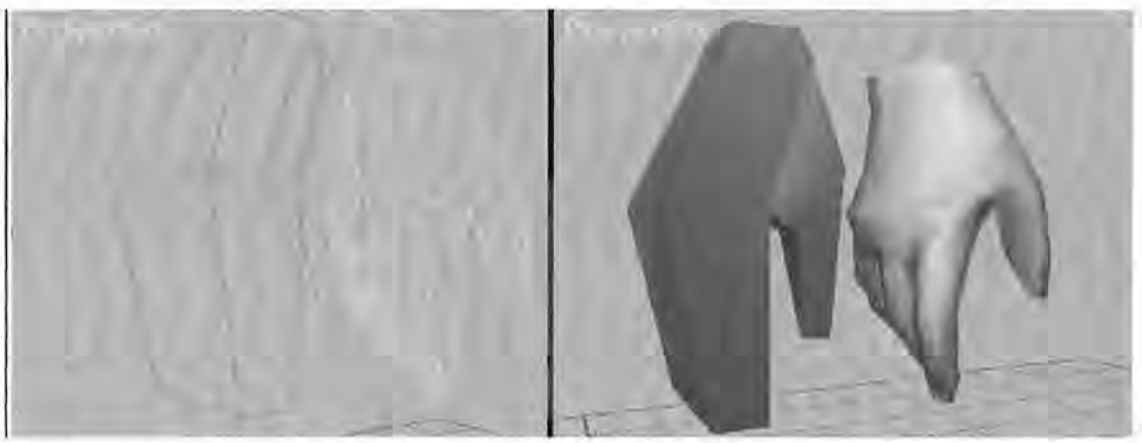

Figure 4.1 The complex hand model (Right) is re-modelled with fewer polygons, which result in a simplified hand model (Left).

In a real social environment, both genders exist and they interact with each other, therefore two different body models must be created, because the male figure is different in representation to the female figure. Instead of create a new set of body parts for the female model, the limbs that are similar in both models are reused (Hands, arms, neck, face and head). We created a dress model to represent the lower body of the female model, while the thighs and legs models are created for the male model. The reason for creating the dress in the female model is that the dress simplified the female model hierarchy and contains fewer polygons than the legs in the male model. In Figure

4.2 and Figure 4.3, the hierarchies of both models are represented in tree graphs; this will illustrate the connectivity between different body parts in both models.

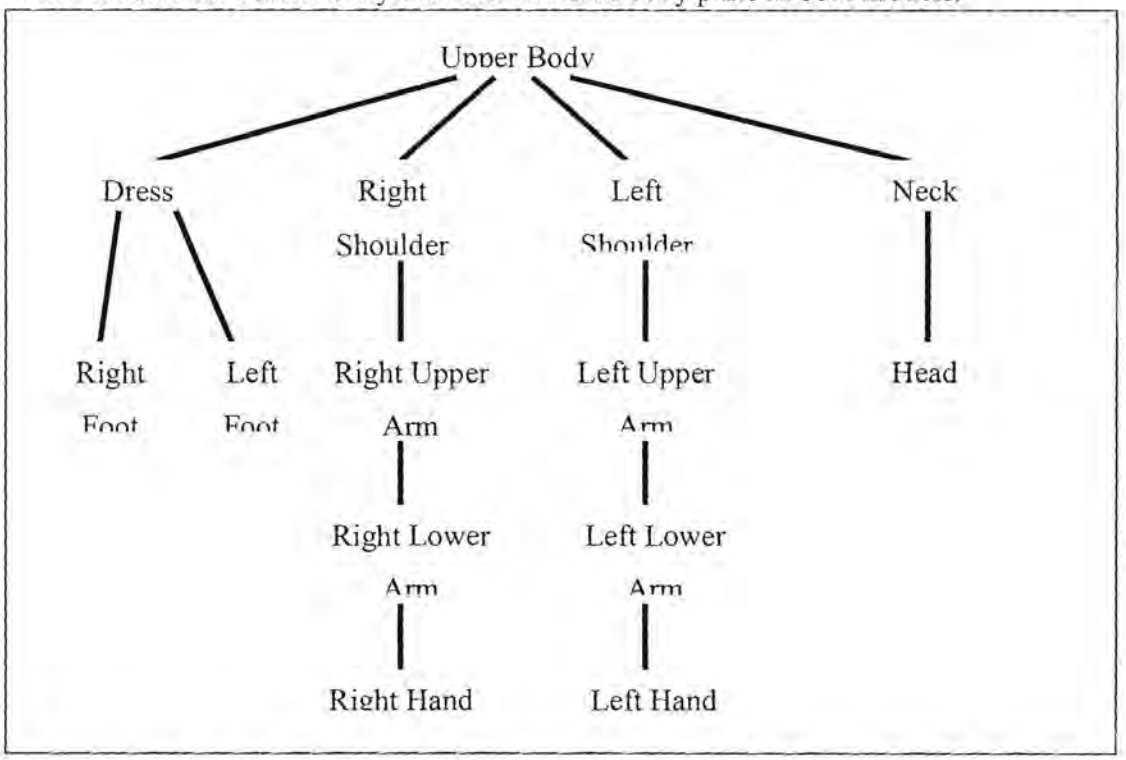

Figure 4.2 Tree hierarchy of the female model

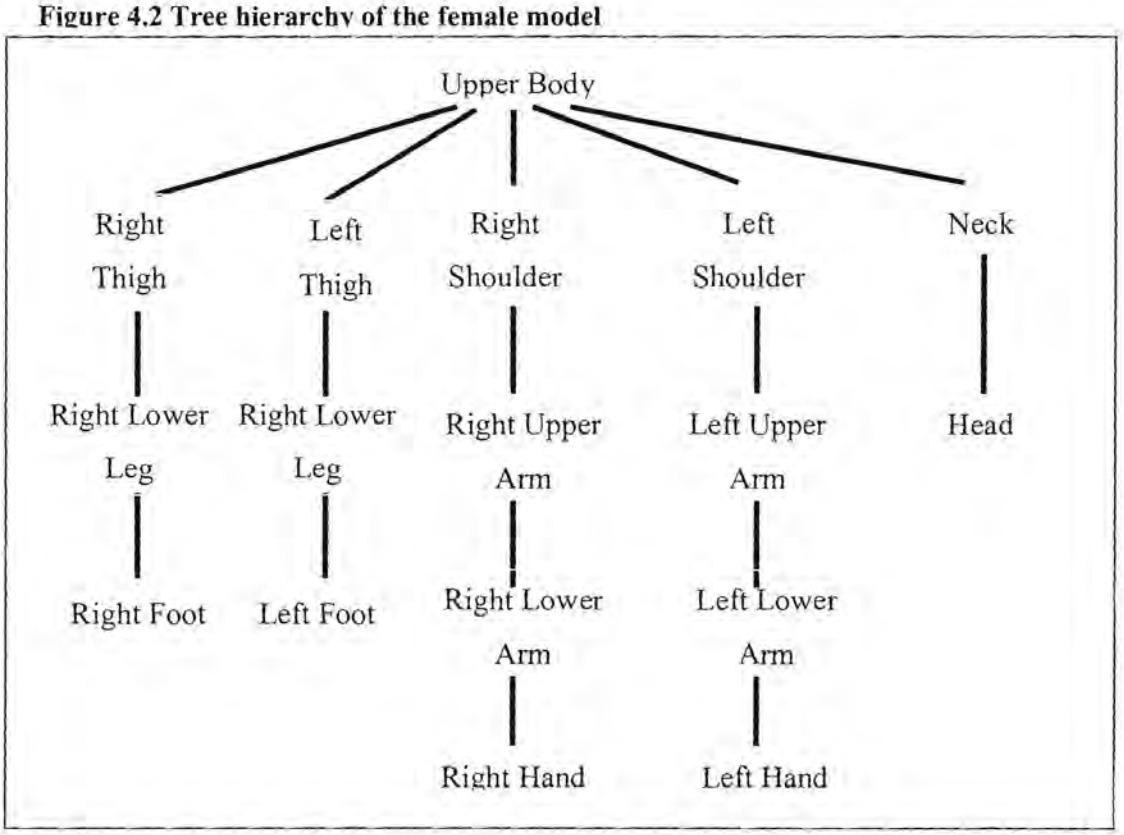

Figure 4.3 Tree hierarchy of the male model

The head model is divided into two parts the face model and the back of the head. The face mesh is reused from expressive texture approach, so those texture mapping, implementation and facial animation do not needed to be re-implemented. The back of the head is modelled such that it fitted behind the existing face mesh. This allows separate texture mapping for the face mesh and the back of the head model.

After the body structure is defined in details, the both body models are model in a 3D-modelling tool before it is implemented in the system. The 3D modelling tool created the prototypes of the body models, which aided the design of body animation for the body models and problems in the models structure are adjusted before implementing the models into the system. E.g. The sizes of different body parts are scaled into correct proportions relative to the body. The prototypes of both body models are shown in Figure 4.4 and Figure 4.5, different colours are allocated to each type of body part, which allows the body part to be identified more easily during body animation.

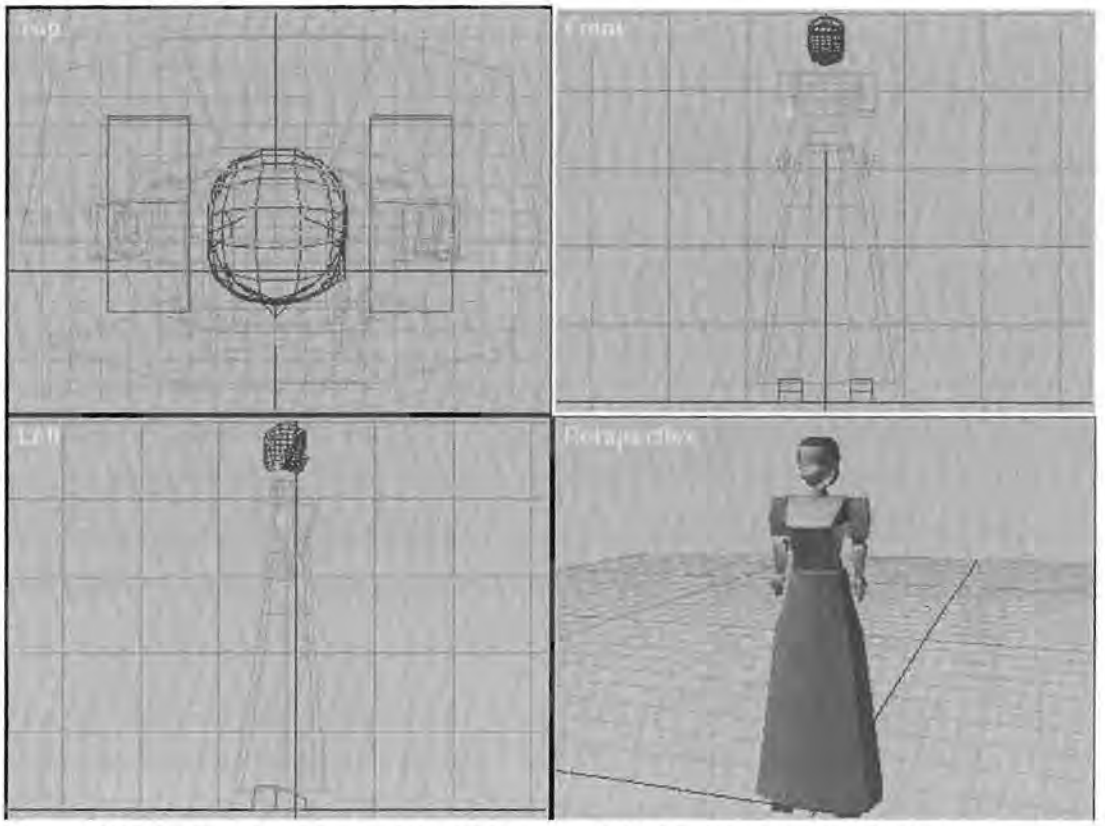

Figure 4.4 The female body model prototype modelled in 3D studio Max R2.5.

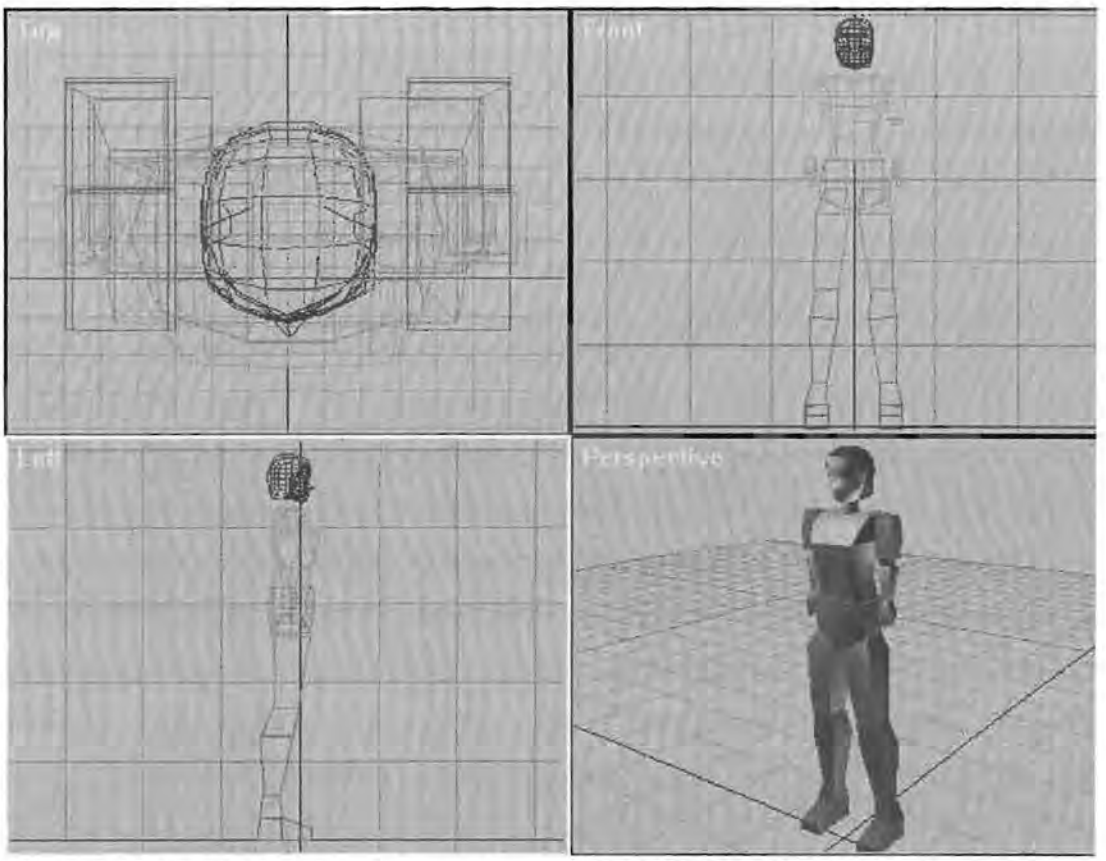

Figure 4.5 The male body model prototype modelled in 3D studio Max R2.5.

The body model prototype allow us to identified body parts that has the same shape and sizes, so during implementation these body parts can be implemented in a modular fashion.

4.1.2. Body Animation

In body animation, the position of all body joint angles must first be determined, and then the motion cues of the body motions must be classified.

Positioning Upper Body Joint Pivots

When animating the body motions, the movement of the avatar body must looked natural to the user, because each joint between the limbs in the human body has angle constraints. E.g. the elbow joint between the lower arm and upper arm of the human body cannot rotate an angle larger than 170 degrees, unless it is broken.

Therefore, the angle constraints and angle positioning in the joints are very important when animating body motions. The angle positioning is determined by the local co-ordinate system of each object model, in many commercial 3D modelling tools, the models created have an initial local co-ordinate system and the object pivot at centre of the object model. This means when the model animates object rotation, it rotates along its pivot axis in the local co-ordinate system centred at the object. When animating the avatar joint rotation the pivot of the body part that requires animating rotation is positioned between two body parts. E.g. When rotating the hand model, the pivot of the hand model is positioned between the hand and lower-arm models (Figure 4.6).

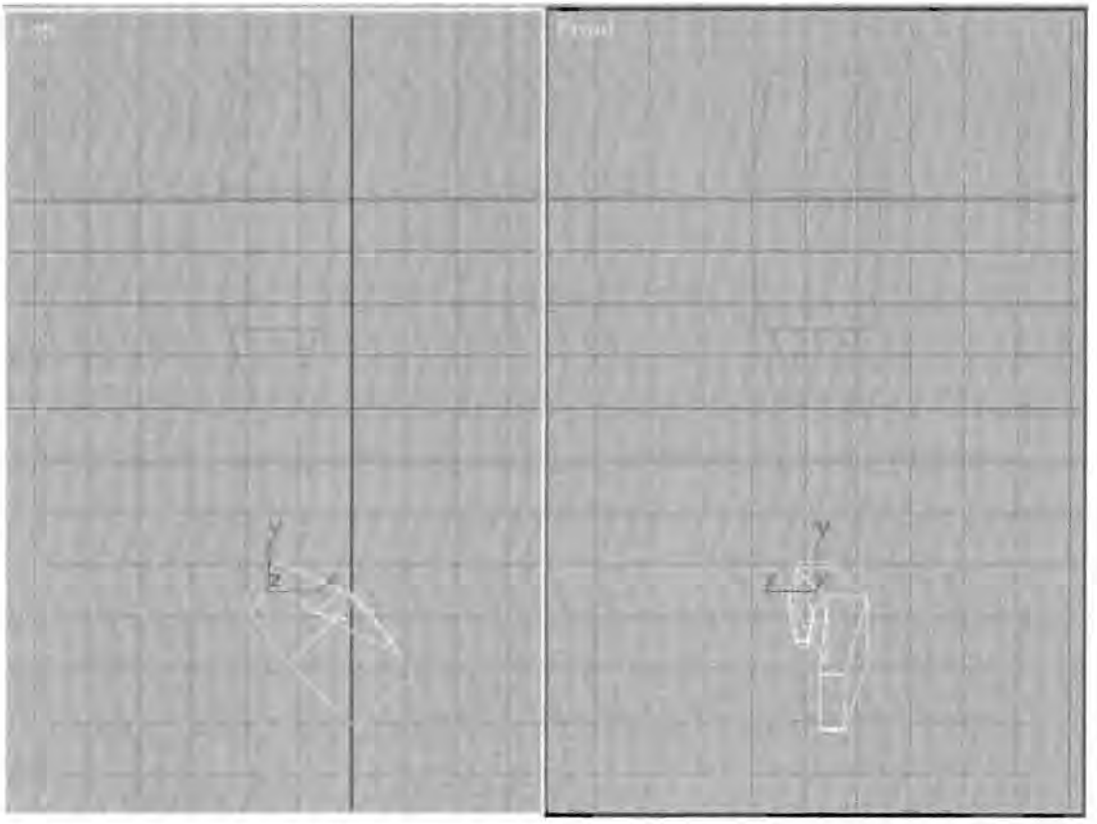

Figure 4.6 The hand model rotation simulating wrist movement.

The body model is structured in hierarchy, this means a consistent behaviour exists between the different parts of the body. The body parts remains connected after rotation and the rotation of the body part at the higher level of the hierarchy will result in the rotation of the lower level object. E.g. If the arm moved, the hand moved with the arm (Figure 4.7), but it the hand moved, the arm remains stilled.

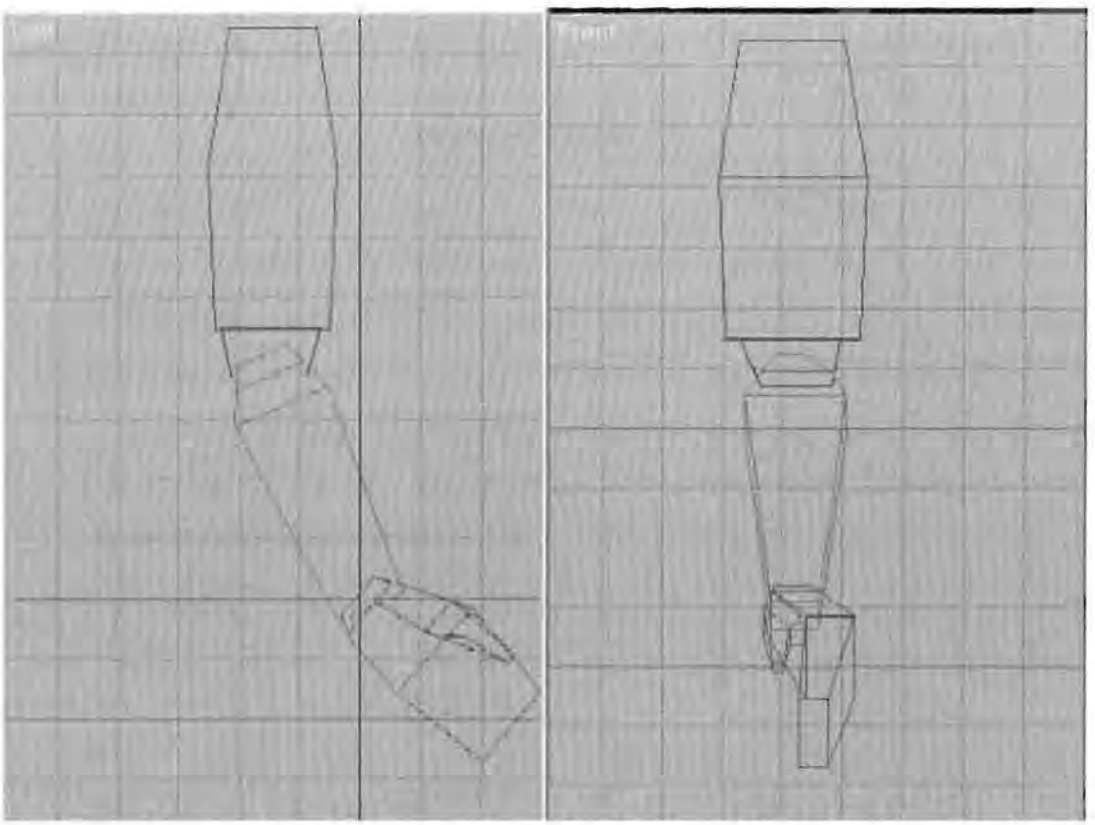

Figure 4.7 The rotation of the arm around its pivot caused the hand moved with the arm, while remain connected with the arm.

Therefore, the pivots of all upper body parts (head, upper arms, lower arms and hand) are adjusted for upper body animation (Figure 4.8).

Body motion cues

After the pivots are defined in the male and female body models, the body motion cues can be classified for each body motion associated with an emotion and action in the environment. During communication and interaction, each individual has unique ways of expressing their emotions using body motions and using gestures. Therefore, the body motion cues classified for each emotion in Table 4.1 is only a basic generalisation. E.g. When Japanese people communicate and interact with each other, they tend to use less or no gestures and body motions, because some body movement are seen as hostile from the Japanese people's point of view. American people on the other hand use gesture and body motions more frequently during communication and interaction.

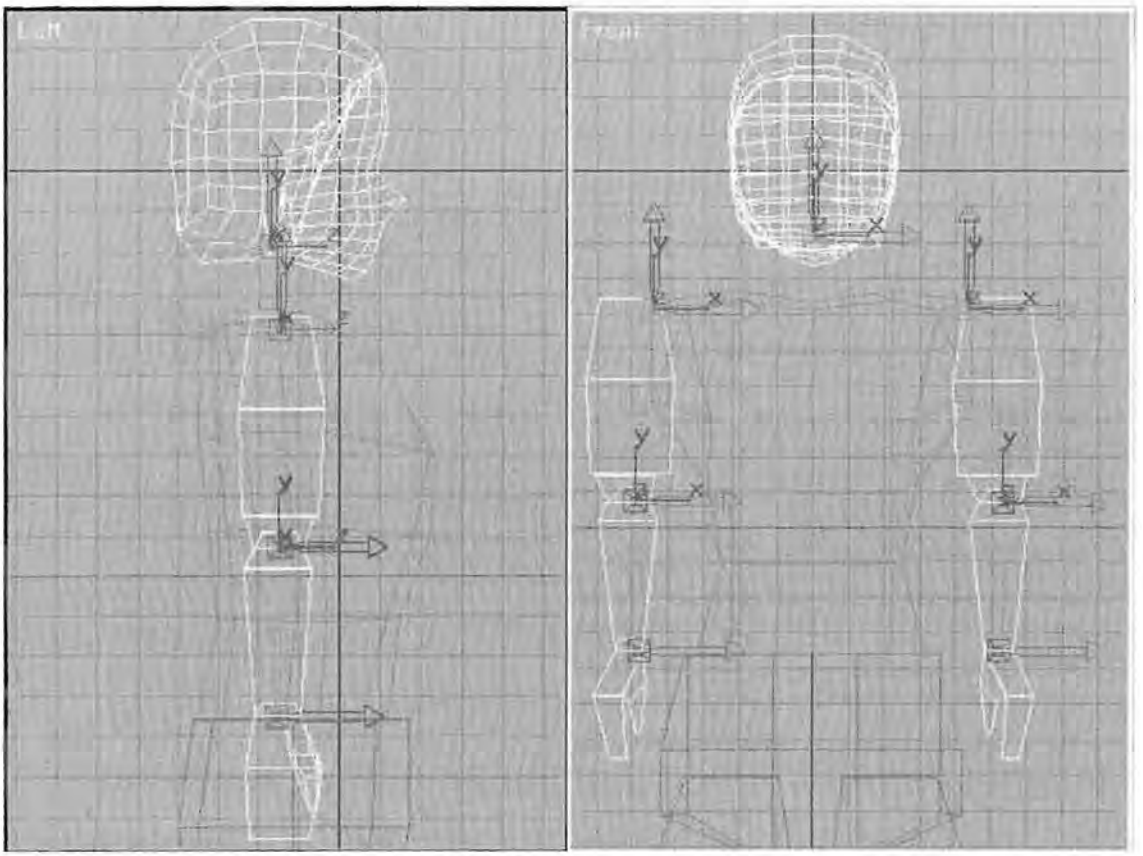

Figure 4.8 The adjusted pivots of all upper body parts for upper body animation.

Emotion	Body Motion Cue
Angry	Head face in the other direction, eyes staring at the respondent, arms crossed.
Melancholy	Head looked down, the hand is placed in front of the face.
Surprise	Head tilted slightly, hand placed over the chest.
Disgust	Head looked slightly away from the respondent, the hand raised in front of the avatar in a "rejection position".
Fear	Head tilted slightly, hand placed against the chest.
Teasing	Head face slightly away from the respondent with eyes focus on the respondent, one arm arched at the waist
Happy	No body movement
Cunning	Head looked in the other direction, arms crossed

Table 4.1 The Body Motion Cues classification for each emotion

The body motions are applied in the same manner for all avatars under a specific emotion and retargetting body motions is not required, because all the avatars are equal in size, this avoids the feeling of superiority from the avatars. From the coven experiment, some users think an avatar larger than a small avatar is more superior. This causes the interaction between avatars that differ in size inflexible [65], because users feel that this is similar to a situation where a subordinate communicates with ones superior.

4.2. Interaction

Before interaction can be determined, we have to create a virtual environment that demonstrates a simple application for the avatars and their interaction in the environment. This allows the users to understand the behaviour of people under a simple, closed social environment without considering the external factors. When avatars interacted in the virtual environment, their actions are dependent on the possible interactions in the environment. Therefore, we decided to create a virtual pub environment for the avatars, because the people communicate and interact with each other in the pub, so the avatars can represent people interacting and communicating in the pub.

In the real environment, interactions are separated into physical interaction and social interactions. Physical interaction, is when people interacting with objects in the environment and Social interaction, is when people interact with other people. Similarly, the avatar interactions in the virtual environment can also be classified into physical and social interactions.

The possible interactions for avatars in the virtual pub environment:

- Drinking (Physical)
- Watch the television in the pub (Physical)
- Change television channel (Physical)
- Talk to other avatars (Social)
- Play music from the juke box (Physical)
- Change music (Physical)

The classification of interactions made it possible to identified which interaction of the avatar is dependent on the objects or other avatars in the virtual environment. E.g. if the avatar talk to other avatars, the interaction is only possible if there are other avatars close by in the virtual environment.

Interaction Restrictions

In the real environment, people cannot interact with all objects and other people, if the object or the other person is not reachable by the individual. Therefore, constraints must be applied to the avatar interactions, so that interaction in the virtual environment is similar to those in the real world environment for the users. Constraints can be applied to interactions defined in the virtual pub environment, so that specific interaction is only possible by the identity of the avatar.

The physical interactions are restricted by object availability, if the object in the environment has limited access then when the object is fully accessed, other avatars cannot interact with the object. E.g. the jukebox can only play one music tune at a time. Therefore, only one avatar at a time can play music from the jukebox. Apart from limited access, the object can also be accessible only by specific avatars. E.g. In the pub, the television is own by the Barman, so only the barman can change the television channel.

The physical interaction restrictions in the virtual pub environment are:

- The television channel is changeable only by the Barman avatar
- The jukebox can be used by any avatar except the Barman avatar, because the barman has to stand at the counter to serve customers.
- Any avatar can change the music on the jukebox
- Any avatar can perform drinking except the Barman
- All avatars can watch television except the Barman

Social interaction is restricted by avatar availability, if there is no avatars surround the avatar wishing to talk then the avatar cannot perform talking with other avatars.

Once all possible interactions for the avatars and constraints are determined, each interaction can be classified into motion cues, which allow us to animate the avatar's body motion during interaction (Figure 4.2).

Interaction	Motion Cue
Drinking	The hand that is holding a drink moved close to the mouth, the hand must return to neutral position after drinking.
Watch the television in the pub	The avatar turn to face the television and look at television with arm cross
Change TV channel	Press the Remote control
Talk to other avatars	Turn to face other surrounding avatar
Play music from the juke box	Move in front of the juke box and switch on juke box
Change music	Move in front of the juke box and switch on juke box

Table 4.2 Interaction Motion Cues

4.3. State Machine

Although the interactions and interaction constraints are defined, but the effects of avatar interactions affecting other avatars and the virtual environment are still not established. In the real world, each action performed by an individual always influenced the emotions of others surround this individual and people interacted with objects for satisfying their goals. If people are prevented or blocked from achieving their goals then people will become frustrated which lead to anger or despair. E.g. If an individual want to used the public phone, but someone is stand blocking the way and not willing to give way, then the individual can become angry.

Therefore, the consequence results of interaction by each avatar will change the emotional state of other avatars and state of the objects in the virtual environment. These emotional and state changes in avatars and objects must be determined and defined in the system as output of the avatar interaction.

The emotional state of an avatar can change because:

- The avatar wants to talk to other avatars, but no other avatars are nearby, so the avatar's emotion changes to melancholy.
- The avatar wants to talk to other avatars, there is other avatar nearby, so the avatar's emotion changed to happy.
- The avatar is watching television, but the Barman changes the channel, therefore the avatar becomes angry with the Barman.
- The avatar is hearing music, and no one is disturbing it, therefore the avatar becomes happy.
- The avatar is hearing music, but another avatar try to change the music tune to another, therefore the avatar is disgust at the other avatar.
- The avatar is shouted by another avatar, therefore the avatar's emotion changes to melancholy.
- The Barman avatar is shouted by another avatar, therefore the Barman avatar's emotion changes to melancholy.
- Result from drinking too much the avatar becomes drunk, which made the avatar looks cunning.

In order to simplify the state changes in objects, only the television and the jukebox in the pub has changeable states with two channels or two music tunes. The state of the television in the virtual pub environment can be showing channel1 or channel2.
The state of the jukebox in the virtual pub environment can be off or on, when it is on it can be playing music tune1 or music tune2.

After all the state changes and interactions are defined, the transitions between interactions and state changes, and interaction processes are implemented into the system using a state machine. A state machine is a finite automation process, which implement the states that an object may be in, as well as the transitions between the states. The state machine is deterministic and the resulting state is unique. The avatars' interactions are synchronised by the state machine, because when an avatar perform an interaction, the interaction is determine in the state machine to identified which virtual

environment object state is changed. These object state changes are stored in the system process and the result of the avatar's interaction is a change in the emotional state of the avatar as the avatar achieved or failed to reach its goal. An avatar can affect other avatar's emotional state by modifying the object's state which cause the other avatar fail to achieve its goals or direct influence (shouting at the other avatar) (Figure 4.9).

The state machine is represented by a state diagram, so that the interactions and transitions in the state machine is scenario tested and corrected before implementation. A state diagram is a model, which describe the states that an object may be in and the transitions between states.

The initial state in the state diagram is the initial Interaction State of the avatar before interacting with the environment and other avatars. The final in the state diagram is the results of the avatar interaction after the avatar interacting with the environment and other avatars. The emotion of the avatar changes or remain the same before interaction.

The rounded rectangle in the state diagram represents the interaction of the avatar, the boxed rectangle represents the objects in the environment and the arrows are representing the events/transitions between the interactions.

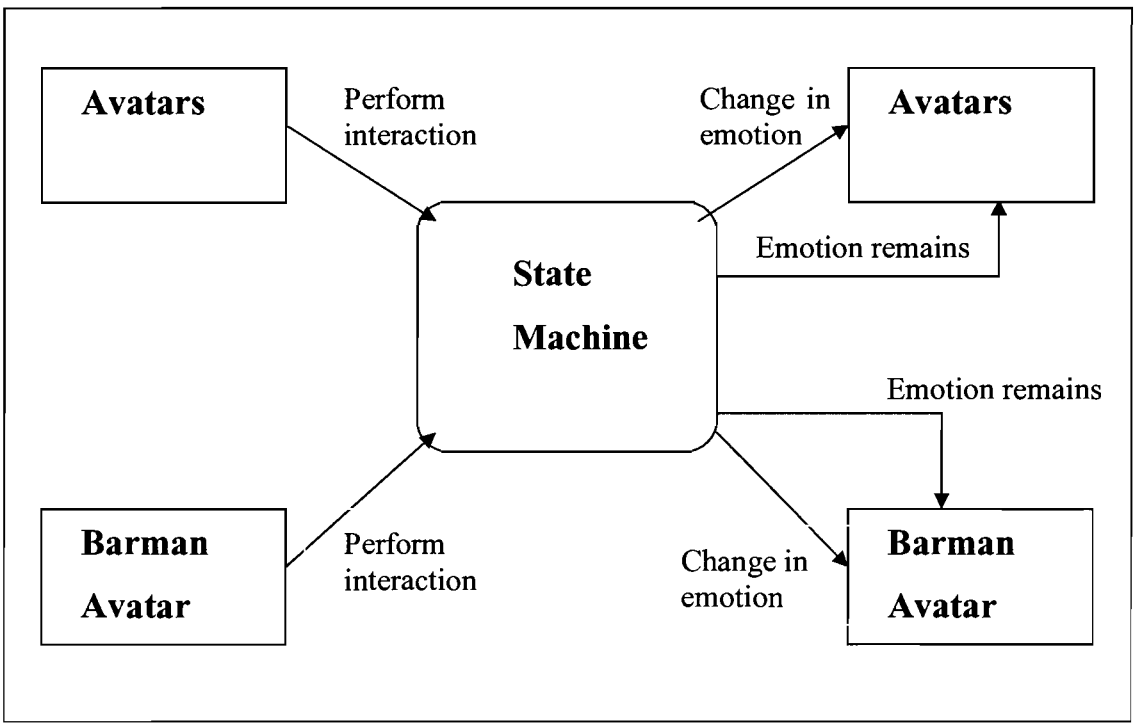

Figure 4.9 The use-case diagram of the state machine.

In Figure 4.10 and Figure 4.11, the state machine is subdivided into different interaction processes, which is easier to understand the different interactions and their connection.

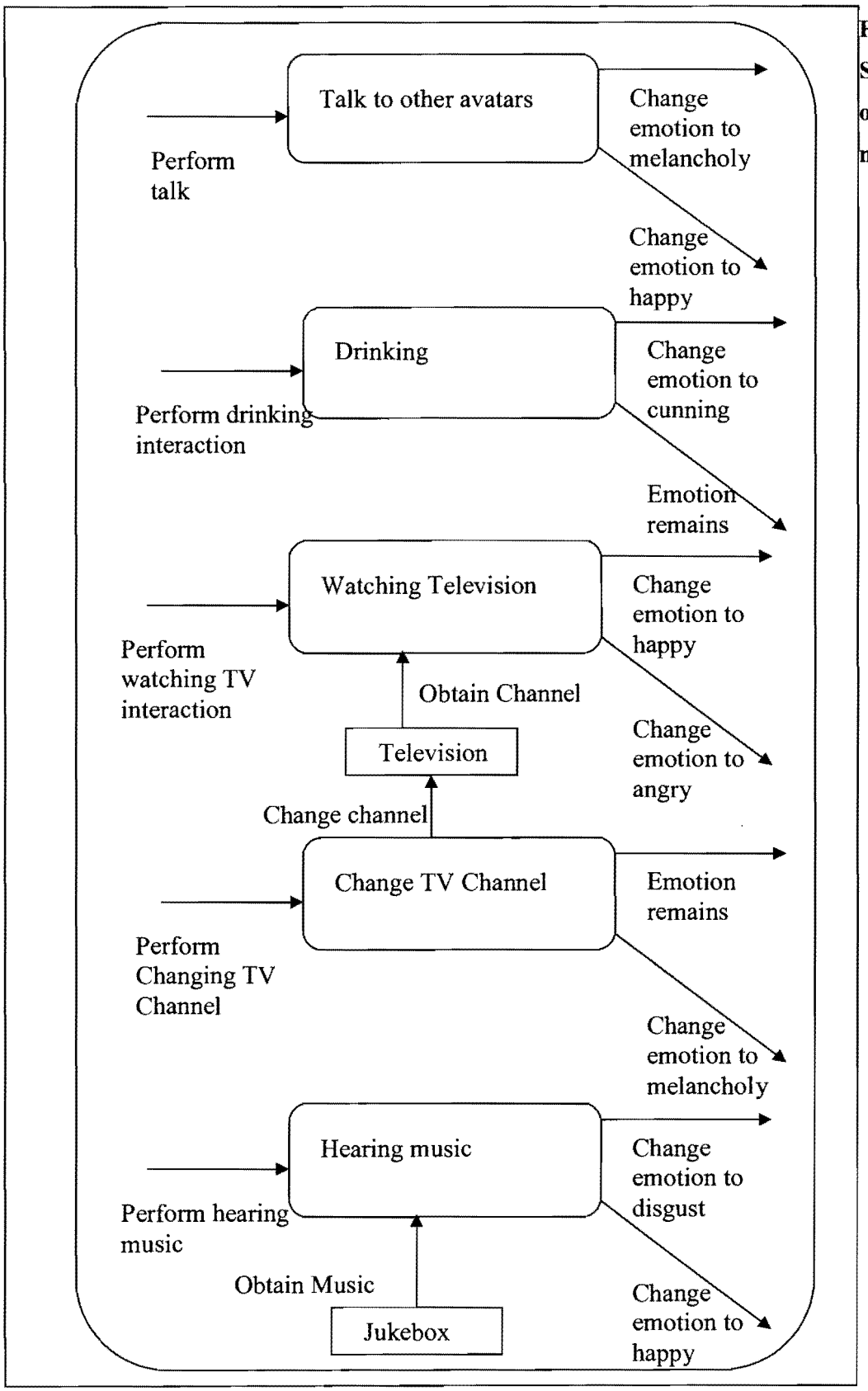

Figure 4.10 State diagram of the State machine.

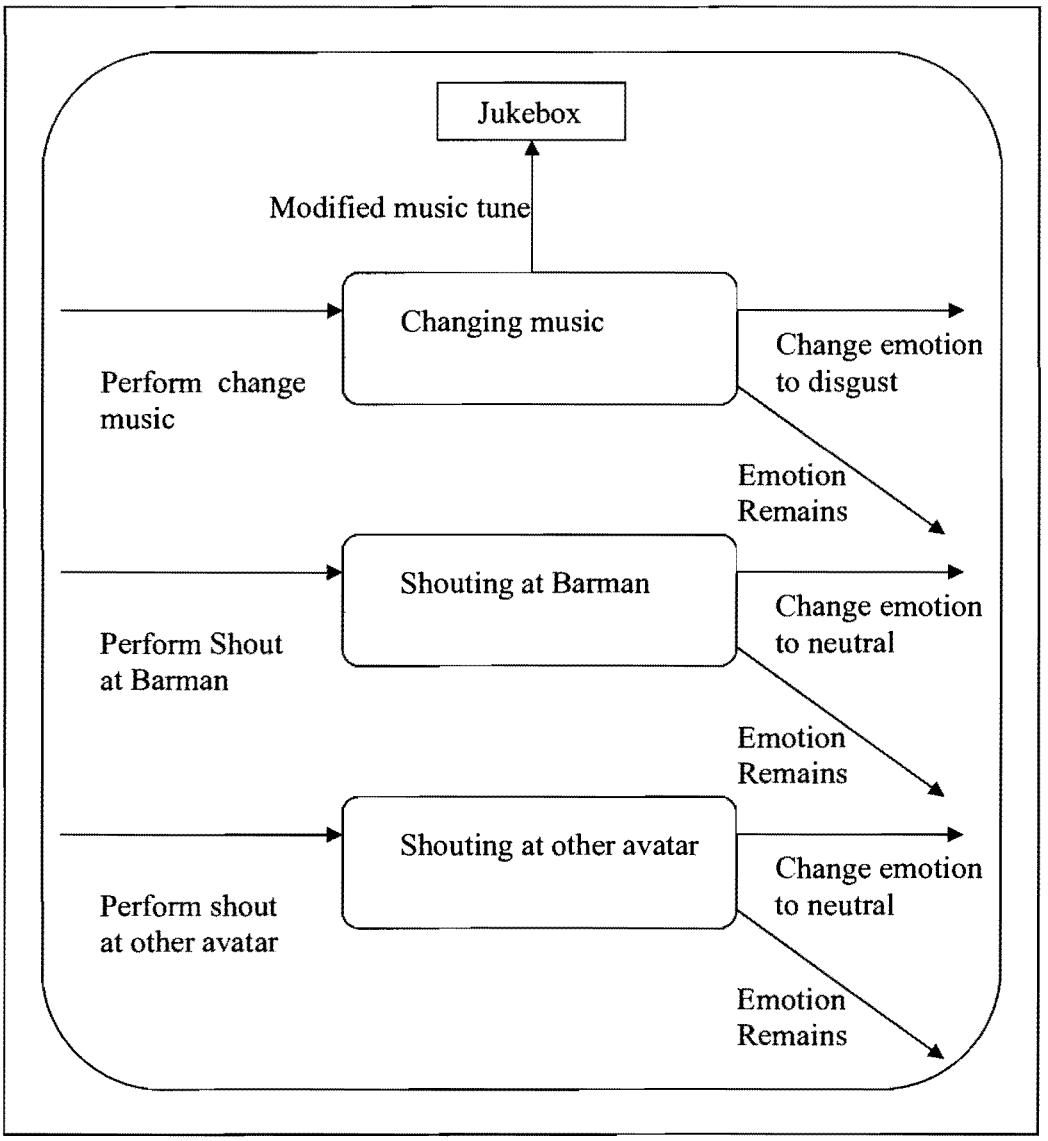

Figure 4.11 (Continue) The state diagram of the state machine.

In figure 4.11, two new interaction processes is added, "shouting at Barman" and "shouting at other avatar". The avatar performed these interaction processes automatically, if the "Hearing music" and "Watching television" interaction process is interrupt by the barman avatar or other avatars. In the real world, people will become angry if another person changed the television while they are watching, and people usually scream at the person who changed the channel to switch back to the previous channel. Therefore, the avatar will continually perform the "shouting at Barman" or

"shouting at other avatar" interaction processes until the avatar that changed the television or jukebox object's state back to previous state. Once the environment objects is returned to previous state, the avatar will continue the interaction process before the interruption by other avatars.

The talking and drinking interactions contained complex internal states that differ to other interaction process. These interaction processes are represented by the internal state diagrams in Figure 4.12 and Figure 4.13, which gives the in-depth view of these interaction processes. In the drinking interaction process, a drinking counter is allocated to each avatar, this counter calculated the number of drinks that each avatar had, if the avatar drink exceed the limit then it is indicated as been drunk.

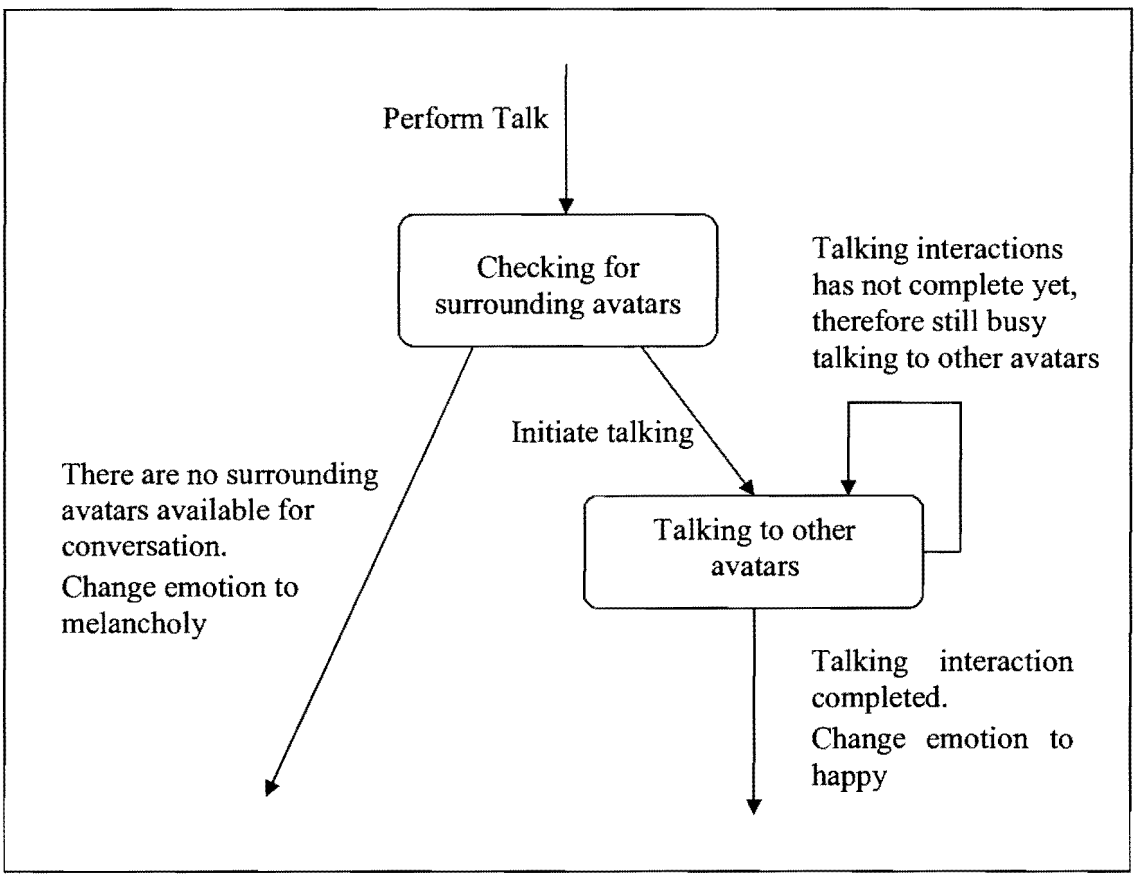

Figure 4.12 Internal state diagram of the talking interaction process.

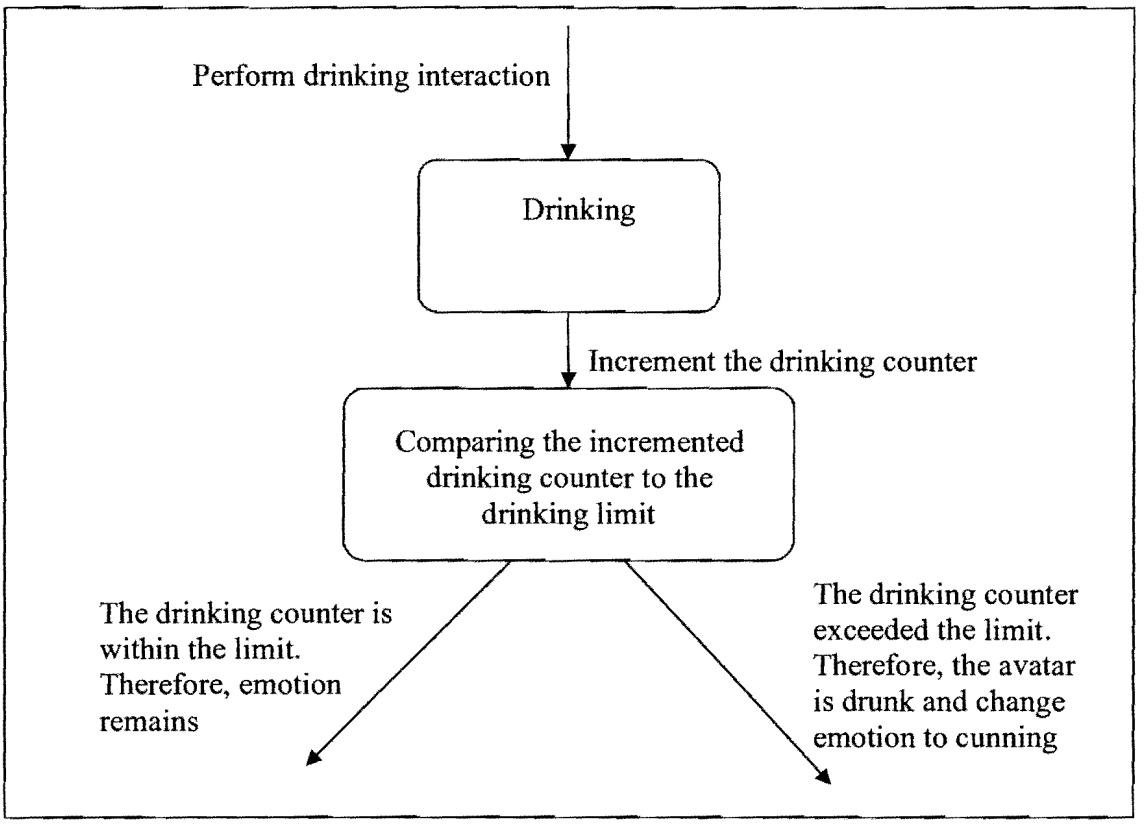

Figure 4.13 Internal state diagram of the drinking interaction process.

The state machine can be implemented according to these state diagrams, but identity checking is required in the actual implementation of the system, before the avatars entering the state machine and perform interactions. This identity checking process determine if an avatar is a customer or the barman and restricted the avatar from performing interactions that it is not supposed to perform.

4.4. Sounds

When avatars interact and communicate in the virtual environment, if sounds are not available, the interaction in the virtual environment will become dull and communication without words is not amusing. In the virtual pub environment, short speech sounds are linked to avatar interactions, so when the avatar perform an interaction, the user can hear the comment made by the avatar. E.g. When the avatar has no one to talk to, the avatar's expression becomes melancholy and it makes a remark

"Boring", then user understand that the avatar is bored because there is no avatars to communicate with it.

Two sets of speech sounds must be provided, because we have to accommodate a set of speech sounds for the male avatar and the other set for the female avatar. The user will found it weird, if the female avatar has a male voice or vice versa. The limitation of this is that all the male or female avatars will have the same voice, because the speech sounds and the environment sounds together required large amount of system resources to loaded. Therefore, the system can only provided one set of speech sounds for each gender. Music tune sounds are also provided in the virtual pub environment when the avatars interacted with the jukebox. This is because the user will not know if the avatar is just standing in front of the jukebox or interacting with it, if music tunes is not played during the interaction processes. The speech and music sounds added more information to the visual interaction made by the avatar in the virtual environment and made interaction more amusing.

4.5. Summary

In order to simulate the physical and social interaction in the virtual environment similar to the real world and study the interactions of the avatars, the avatar body structure must be modelled such that animating interactions is possible. When modelling the avatar, gender must also be taken into account, because both genders exist in the real environment and they differ in appearance. Since the male and female models has equal size, and similar upper body structure, the body motion of the male avatar can be applied to the female avatar or vice versa without retargetting. By presenting interactions of the avatar in state diagram, the transitions between interactions are easier to understand and error appeared in the transition between interactions are detected and corrected before implementation.

The next chapter will discuss the actual design and implementation of the virtual pub application and the texture mapping avatar faces tool, which takes the theoretical approaches and prototypes into implementation.