Simulation of Small Group Discussions for Middle Level of Detail Crowds

Jigish Patel Robert Parker David Traum
Institute for Creative Technologies, University of Southern California
13274 Fiji Way, Marina del Rey, CA 90292 USA
traum@ict.usc.edu

Abstract

We present an algorithm for animating middle level of detail crowds engaged in conversation. Based on previous work from Padilha and Carletta, this algorithm is used to provide gestures for group characters in an embedded virtual world. The algorithm is implemented and used within the Mission Rehearsal Exercise project at ICT to control Bosnian crowd members.

1 Introduction

In the Mission Rehearsal Exercise at University of Southern California, (Swartout et al., 2001), a leader is trained with a story-based immersive simulation including many characters, both teammates and others, as shown in Figure 1. While a number of these characters (especially the ones in the front and center) play lead roles in the story and interact heavily with the trainee (Rickel et al., 2002), there are also a number of "supporting" characters who play fairly minor roles, but are still important to the setting of the story, e.g., the crowd members shown to the far right. The original versions of these characters had all their motions painstakingly hand animated, and were set in loops when the interaction lasted longer than the amount of scripting. Such scripting has three problems: first it is labor intensive, second, it is not reactive to local circumstances, and third, the repetition can detract from the realism, even if well animated for short segments.

A solution to these problems is to use some automatic simulation rather than hand-scripting. As (O'Sullivan et al., 2002) point out, crowd and group simulations are becoming increasingly important for a number of applications, including movies, as well as games and simulations. Random or scripted behaviors are satisfactory for low levels of details (e.g., very distant crowds), and full animated conversational agents are adequate for the main characters, but these are overkill for middle-level group members who are seen at some distance and not directly interacted with. What we need for our middle-level characters is something good enough to look like characters involved in conversation without the overhead of fully intelligent agents. A very good starting point is provided by (Padilha and Carletta, 2002; Padilha and Carletta, 2003), who synthesize some of the best research on group dialogue behavior into a parameterizable, probabilistic algorithm for individual behavior as part of a group. We have re-implemented this simulation, with some enhancements, and used the results to animate the Bosnian crowd members in the Mission Rehearsal Exercise.

2 Animating Characters using Crowd Simulation

The main focus of (Padilha and Carletta, 2002) is a simulation of the external view of a group conversation. The group is assumed to be part of a conversation, taking turns speaking, and listening to others. Much of the literature on turn-taking is synthesized at a fairly abstract level, including the pacing of turn-taking, listener behavior and feedback, and physical behaviors such as gaze, head movement, gesture, and posture shifts. The characters are not given any actual content to speak or more detailed topics of conversation. For middle-level of detail characters, not in the foreground, this is not necessary in any case - an observer is not meant to listen to the conversation or follow it closely.

Padilha and Carletta consider several basic behaviors, including:

- speech (including beginning or continuing or concluding a turn, or giving positive or negative feedback)
- head/face nods and facial expressions
- gestures
- posture shifts
- gaze (at other characters)
- listening

Behaviors are generated by simulating a probabilistic algorithm. There are a number of parameters (see below) which can take on values between 0 and 1. Each character has a value for each parameter, and runs in loop, testing random values against the parameter, and deciding which behaviors to do based on the results (as well as external conditions, such as what other characters are doing. Although (Padilha and Carletta, 2002) have an implemented simulation algorithm with results specifying behaviors such as the above, they did not actually link up the simulation to an animation system. We have linked up with BDI's Peopleshop^{TM} characters, using provided basic animations for those characters. This necessitated making individual choices of

Figure 1: The MRE domain: Bosnian Group shown at right.

types of gestures to indicate speaking and other motions. Figure 2 shows a snapshot of the characters involved in conversation.

Figure 2: Bosnian Group in conversation

We have also made several extensions to the simulation of (Padilha and Carletta, 2002) to account for the use of this simulation as embedded in the virtual world. First some extensions to the gaze model, to account for change of addressee and audience gaze at multiple speakers (since a conversation may sometimes fragment into smaller group conversations among sets of the participants. More importantly, though, we also allow attention to pass away from the group discussion to focus on external events such as the main conversation between the human trainee and main character virtual humans and other occurrences, such as explosions and people and vehicle movements.

These parameters were defined in (Padilha and Carletta, 2002):

talkativeness: likelihood of wanting to talk.

transparency: likelihood of producing explicit positive and negative feedback, and turn-claiming signals.

confidence: likelihood of interrupting and continuing to speak during simultaneous talk.

interactivity: the mean length of turn segments between TRPs.

verbosity: likelihood of continuing the turn after a TRP at which no one is self-selected.

In addition, we added the following parameters:

responsiveness: likelihood of a participant reacting to interruptions from outside the group.

continuity: likelihood of selecting an addressee (for example, by asking a question to him/her specifically) at the end of the speaker's turn.

The parameters can be changed for different runs. Figure 3 shows the control panel to change the parameters (one character, Serge, is highlighted, but others can be selected as well by clicking on their name on the left). One can select each agent and configure the parameters independently, thus changing the nature of the conversation. One can give individual characters different "personalities" by configuring the parameters differently. For example, a character would talk more with high talkativeness and high verbosity. A confident character would have high confidence. An active character would have both high transparency and high talkativeness. A bored character would have low verbosity and tranparency and high responsiveness to external events.

3 Algorithm

The simulation runs by cyclically testing the set of parameter values against random numbers, with the results leading to decisions of whether to speak or listen or attend elsewhere and which gestures to make. A loop of the algorithm in Figure 4 is executed every cycle (approximately 500 ms long) by each character. This is a

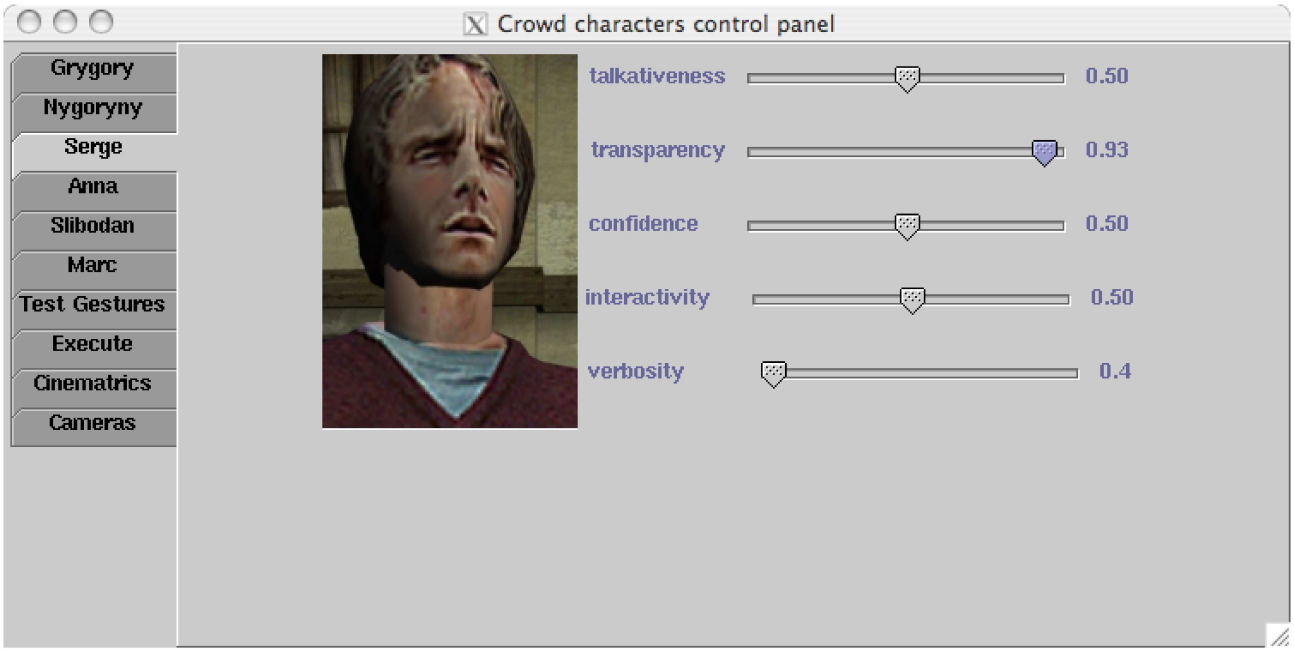

Figure 3: Control Panel for Minor Character Behavior settings

modification of the algorithm in (Padilha and Carletta, 2002). The main modifications involve allowing agent responsiveness to events and speech outside the group and the linking of abstract behaviors to specific animation calls for the characters. The main conditions of the algorithm concern who was previously speaking. Conditions include:

- no one
- self
- someone else being listened too
- self and someone else (further divided into conditions of self speaking first, new simultaneous start, and continued simultaneous speech)

Within each major segment, the agent must decide whether to (continue to) speak, whether to listen, and what kind of feedback and gestures to perform. Changes to the turn status will then affect future iterations.

4 Evaluation

Padilha and Carletta's evaluation plan involved comparing their simulation to transcripts of group conversation data, showing a better fit than simpler models. While this kind of evaluation would certainly be interesting, we propose a different kind of evaluation - whether the simulation "looks like a conversation" to a viewer. Two baselines for performance are whether the simulation looks more natural than random motion and whether the simulation looks more natural than the looping, hand-crafted animation. We have implemented the algorithms with behaviors in the MRE Bosnia scenario, both as a stand-alone demo, and as part of the general training scenario.

We also want to evaluate the effects of the individual parameters. We have constructed experiments in which different characters are given different values for parameters (such as talkativeness and confidence), and then showed viewers recordings of different simulation runs with these parameters to judge features like apparent talkativeness of individual characters.

References

O'Sullivan, C., Cassell, J., Vilhjalmsson, H., Dingliana, J., Dobbyn, S., McNamee, B., Peters, C., and Giang, T. (2002). Levels of detail for crowds and groups. *Computer Graphics Forum*, 21(4).

Padilha, E. G. and Carletta, J. C. (2002). A simulation of small group discussion. In *Proceedings of EDILOG 2002: Sixth Workshop on the Semantics and Pragmatics of Dialogue*, pages 117–124.

Padilha, E. G. and Carletta, J. C. (2003). Nonverbal behaviours improving a simulation of small group discussion. In *Proceedings of the First International Nordic Symposium of Multi-modal Communication.*

Rickel, J., Marsella, S., Gratch, J., Hill, R., Traum, D., and Swartout, W. (2002). Toward a new generation of virtual humans for interactive experiences. *IEEE Intelligent Systems*, 17.

Swartout, W., Hill, R., Gratch, J., Johnson, W., Kyriakakis, C., Labore, K., Lindheim, R., Marsella, S., Miraglia, D., Moore, B., Morie, J., Rickel, J., Thiebaux, M., Tuch, L., Whitney, R., and Douglas, J. (2001). Toward the holodeck: Integrating graphics, sound, character and story. In *Proceedings of 5th International Conference on Autonomous Agents.*

```
If(interrupted)
        Test responsiveness to react to the interruption
If (no one is speaking)
        test talkativeness to start to speak here
        if so,start with a random interval
              test transparency to shift posture.
If (listening to a single speaker)
        look at him; occasionally, look away
        if (read the pre-TRP)
              test talkativeness to decide to start
              if so,test transparency
                    to make turn-claiming signals now
        mark next cycle as the TRP
        if (at a TRP and decided to start)
        or (at a TRP and was selected)
              start with a random interval
              test transparency to shift posture
        if (at a free TRP and not going to start)
              test transparency to do feedback
        if (anywhere else)
              test talkativeness and confidence
              to start to speak, i.e.  to interrupt.
If (started simultaneously at a TRP)
        Reduce confidence depending on who started first
        test confidence to decide whether to continue.
        if ( speaking )
              select an addressee at random
              gaze at the selected participant
        else
              select one of the speakers at random with higher probability of selecting
              the one who starts first;
              gaze at the selected speaker
If (speaking simultaneously, and not planning)
        test confidence whether to continue.
        if ( speaking )
              select an addressee at random
              gaze at the selected participant
        else
              select one of the speakers at random with higher probability of selecting
              the one who starts first;
              gaze at the selected speaker
If (speaking alone in a turn)
        use interactivity to set the segment length
        decide when to gesture and gaze-away
        decide when planning stage ends
              gaze back at interlocutor at that point
        if (at the last cycle before the TRP)
              test continuity to select next-speaker
        if (arrived at a TRP and no one started)
              test verbosity to continue talking.
```

Figure 4: Conversational Character Algorithm