Dynamic Movement and Positioning of Embodied Agents in Multiparty Conversations

Dusan Jan
USC Institute for Creative Technologies
13274 Fiji Way
Marina del Rey, CA 90292
jan@ict.usc.edu

David R. Traum
USC Institute for Creative Technologies
13274 Fiji Way
Marina del Rey, CA 90292
traum@ict.usc.edu

ABSTRACT

For embodied agents to engage in realistic multiparty conversation, they must stand in appropriate places with respect to other agents and the environment. When these factors change, such as an agent joining the conversation, the agents must dynamically move to a new location and/or orientation to accommodate. This paper presents an algorithm for simulating movement of agents based on observed human behavior using techniques developed for pedestrian movement in crowd simulations. We extend a previous group conversation simulation to include an agent motion algorithm. We examine several test cases and show how the simulation generates results that mirror real-life conversation settings.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence; I.6.3 [**Simulation and Modeling**]: Applications

General Terms

Algorithms,Experimentation

Keywords

embodied conversational agents, proxemics, multiparty conversation, social interaction, simulation

1. INTRODUCTION

When we look at human conversation in a casual, open setting, such as a party or marketplace, one of the first things we notice is a tendency for people to cluster into sub-groups involved in different conversations. These groupings are not fixed, however, people will often join and leave groups and often move from one group to another. Groups themselves may fragment into subgroups, and smaller groups sometimes merge into one larger group. Participants in these groups adapt their positions and orientations to account for these circumstances, often without missing a beat or otherwise disrupting their conversations.

In order to create believable social environments for games or training simulations we need agents that can perform these same kinds of behaviors in a realistic way. Crowd simulations place an emphasis on large scale movement [9, 3] while work on embodied conversational agents focuses on aspects such as gaze, facial expressions, and hand and arm gestures [2]. There is some important work on authored presentation agents and avatars for human participants which take account of position in the modelling (e.g. [10, 7]), but none of this work presents fully explicit algorithms for controlling the positioning and movement behavior of autonomous agents in dynamic conversations.

In previous work, it has been shown that incorrect positioning of animated agents has a negative effect on believability of dynamic group conversation [4]. Research from anthropologists and social psychologists such as classic work on proxemics by Hall and positioning by Kendon [5] explain social reasons for how people position themselves in different situations. It is also important to know that people expect similar behavior in virtual environments as in real life as shown in [1]. This gives us basic principles on which to base the simulation and provides some qualitative expectations, but is not suitable to directly convert into algorithms. The social force model [3] developed for crowd simulations gives a good framework for movement simulation. While the basic model shows how to handle pedestrian motion we apply the model to the problem of movement in conversation setting.

The basis for our work is simulation of small group discussion as described in [4, 6]. This simulation allowed dynamic creation, splitting, joining, entry and exit of subconversations. However, the characters were located in fixed positions. As indicated in subject evaluations, this significantly decreased believability when conversation groups did not coincide with positioning of the agents. Since the main goal was to maintain believability, adding support for movement of characters is a natural step to counter these less believable situations. We augment the work of [4] by adding a movement and positioning component that allows agents to monitor "forces" that make it more desirable to move to one place or another, iteratively select new destinations and move while remaining engaged in conversations.

2. REASONS FOR MOVEMENT

There are several reasons why someone engaged in conversation would want to shift position. The first reason we consider for repositioning of conversation participants is au-

AAMAS'07 May 14–18 2007, Honolulu, Hawai'i, USA.

978-81-904262-7-5 (RPS) ©2007 IFAAMAS

dibility of speaker. This can come from two factors, either the absolute volume of the speaker, or the relative volume compared to other "noise". Noise here is meant to describe all audio input that is not speech by someone in the current conversation group. This includes speech of agents engaged in other conversations as well as non-speech sounds. When we are comparing loudness of different sources we take into account that intensity of perceived signal decreases with the square of the distance and also that loudness of several sources is additive.

Even when the speaker can be heard over a noise source, if outside disruptions are loud enough, the group might want to distance itself into a more remote area where they can interact without interruptions. Each of the participants may decide to shift away from a noise source, even without an explicit group decision. Of course this may not always be possible if the area is very crowded.

Another reason for movement is proxemics. Individuals generally divide their personal space into four distinct zones. The intimate zone is used for embracing or whispering, the personal zone is used for conversation among good friends, the social zone is used for conversation among acquaintances and the public zone for public speaking. The actual distances the zones span are different for each culture and its interpretation may vary based on an individual's personality. If the speaker is outside the participant's preferred zone, the participant will move toward the speaker. Similarly if someone invades the personal zone of a participant, the participant will move away.

The final reason for movement is specific to multiparty conversations. When there are several people in conversation they will tend to form a circular formation. This gives the sense of inclusion to participants and gives them better view of each other [5].

3. SOCIAL FORCE MODEL

We present our movement simulation in the context of a social force model. Similar to movement in crowds, movement of people while engaged in conversation is to a large extent reactionary. The reaction is usually automatic and determined by a person's experience, rather than planned for. It is possible to assign a vectorial quantity for each person in conversation, that describes the desired movement direction. This quantity can be interpreted as a social force. This force represents the influence of the environment on the behavior of conversation participant. It is important to note however that this force does not directly cause the body to move, but rather provides a motivation to move.

We associate a force with each reason for movement:

$\vec{F}_{speaker}$: attractive force toward a speaker

$\vec{F}_{noise}$: repelling force from outside noise

$\vec{F}_{proximity}$: repelling force from agents that are too close

$\vec{F}_{circle}$: force toward circular formation of all conversation participants

$\vec{F}_{speaker}$ is a force that gets activated when the speaker is too far from the listener. This can happen for one of two reasons. Either the speaker is not loud enough and the listener has to move closer in order to understand him, or he is outside the desired zone for communication. When the agent decides to join a conversation this is the main influence that guides the agent to his conversation group.

$\vec{F}_{noise}$ is a sum of forces away from each source of noise. Each component force has direction away from that particular source and its size is inversely proportional to square of the distance. This means that only sources relatively close to the agent will have significant influence. Not all noise is a big enough motivation for the agent to act upon. The force is only active when the noise level exceeds a threshold or when its relative value compared to speaker level in group exceeds a threshold.

$\vec{F}_{proximity}$ is also a cumulative force. It is a sum of forces away from each agent that is too close. The force gets stronger the closer the invading agent gets. This takes effect for both agents in the conversation group and other agents. This is the second force that is modeling proxemics. While $\vec{F}_{speaker}$ gets activated when the agent is farther than the desired social zone, $\vec{F}_{proximity}$ is activated when the agent moves to a closer zone. Based on how well the agents know each other this can be either when the agent enters intimate zone or personal zone.

$\vec{F}_{circle}$ is responsible for forming the conversational group into a convex, roughly circular formation. Each agent has a belief about who is currently participating in the conversation. An agent will compute the center of mass of all these assumed participants and average distance from the center. If an agent's position deviates too much from the average, the $\vec{F}_{circle}$ gets activated either toward or away from center of mass. Notice that $\vec{F}_{proximity}$ takes care of spreading out around the circle.

As described above, each force has some conditions that determine whether the force plays an active role in motivating movement. Since the forces are not actually physically acting on agent's bodies, it is not unreasonable for agents to suppress a certain force. All the possible causes for movement are always present, but the agents selectively decide which ones they will act upon in a given situation. This is unlike a kinematics calculation with physical forces where all forces are always active. Combining all the conditions we can define which forces are active according to a simple decision procedure. We can view this as priorities the agent has that decide which conditions are more important to react to.

In our implementation we use the following priorities. If the speaker is not loud enough only $\vec{F}_{speaker}$ and $\vec{F}_{proximity}$ are active. If noise is louder than the speaker then $\vec{F}_{noise}$ is added to the mix. If noise is too loud then only $\vec{F}_{noise}$ and $\vec{F}_{proximity}$ are active. If all sound level conditions are met then $\vec{F}_{proximity}$ is active if someone is too close, otherwise we have just $\vec{F}_{circle}$.

Using the above priorities we have a force defined at each point in space where an agent could be located. We do not use this for continuous computation of movement, but rather use it to compute destination points. In each planning cycle the agents will consider whether they should move. To do this an agent considers his position in the force field and computes a destination in the direction of the force field. This process is performed iteratively a constant *bound* times (unless there is no movement in an earlier iteration).

Once we have computed the destination, we use it as a destination point for the character movement algorithms in the Unreal Tournament game engine. These will manage character animation and collision avoidance.

4. TEST CASE ANALYSIS

A full evaluation of the social-force based positioning algorithm presented in the previous section would involve analysis of simulations to see if they improve believability over static simulations such as [4], or other algorithms. While this remains future work for the moment, we did evaluate the algorithms against a series of test cases where we know what behavior to expect from known forces.

In the simulations we describe here we did not change the conversational attributes of agents, but we did constrain the grouping dynamics. In a normal situation the agents would randomly form conversation groups, based on their stochastic decisions. Here we wanted to examine particular scenarios and how the movement algorithm would react to specific changes in conversation group structure. For this reason we disabled conversational grouping decisions in the algorithm and triggered the group structure changes manually from the user interface.

The only variable input to the movement algorithms for different agents is the preferences for proxemics. Each agent has defined values for all zones, but we set all agents to use social zones for communicating. The other parameters such as thresholds for hearing a speaker and noise and circular formations were fixed for these experiments.

In the first case we had an agent approach a group of three agents that were involved in conversation. As the new agent engaged the others correctly spread out for the new agent to occupy the created space and join the formation. In second test case 6 agents were involved in discussion and 2 of them start a side conversation. As agents disambiguate the situation from interruption they correctly distance away from each other as they form two conversation groups. In the last test we examined the behavior that sometimes happens in conversations of people from different cultural backgrounds. Two agents are involved in conversation, but their social zones are incompatible. The result is that one agent continues to adjust his position closer to the other agent while the other continues to back away as observed in [8]. In all of these examples, the simulation generated expected behavior for the agents.

5. CONCLUSIONS

In the previous section, we have shown examples of how the movement algorithm can mirror many effects we see in real conversations. The examples however were very constrained and could not show all the possible combinations that could result from random choices the agents can make. Given the fact that each agent maintains his own belief about who is currently in their conversation we can see many interesting effects when those beliefs become unsynchronized.

As seen in the third test case, we can get some very interesting results when we simulate agents of different cultures. We think that this simulation approach can be fruitful for modeling cultural differences in conversational behavior, and could be used for inter-cultural and cross-cultural awareness and training. We are currently exploring whether we can model different cultural norms for conversational behaviors in ways such that the resulting agent interaction can be recognized as appropriate to one culture or another.

There are still several improvements possible for the conversation simulation. On the presentation side we are planning to make some improvements to the bodies and number and types of conversational gestures they can display. We also plan to improve the algorithm so that it will be able to generate different conversation styles. Currently all conversations take the same form where all the agents have the same goals, their only goal is to engage in conversation with other agents. We plan to introduce the notion of tasks so that we can better simulate different kinds of activities such as asking for directions, a political debate, or casual conversation.

6. ACKNOWLEDGMENTS

The project described here has been sponsored by the U.S. Army Research, Development, and Engineering Command (RDECOM). Statements and opinions expressed do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

7. REFERENCES

[1] J. N. Bailenson, J. Blascovich, A. C. Beall, and J. M. Loomis. Interpersonal distance in immersive virtual environments. *Personality and Social Psychology Bulletin*, 29:819–833, 2003.

[2] J. Cassell, J. Sullivan, S. Prevost, and E. Churchill, editors. *Embodied Conversational Agents*. MIT Press, Cambridge, MA, 2000.

[3] D. Helbing and P. Molnár. Social force model for pedestrian dynamics. *Phys. Rev. E*, 51(5):4282–4286, May 1995.

[4] D. Jan and D. R. Traum. Dialog simulation for background characters. *Lecture Notes in Computer Science*, pages 65–74, 2005.

[5] A. Kendon. *Spatial Organization in Social Encounters: the F-formation System*, pages 209–237. Cambridge University Press, 1990.

[6] E. Padilha and J. Carletta. A simulation of small group discussion. *Proceedings of EDILOG 2002: Sixth Workshop on the Semantics and Pragmatics of Dialogue*, pages 117–124, 2002.

[7] M. Rehm, E. Andre, and M. Nischt. Let's come together - social navigation behaviors of virtual and real humans. In M. M. et al., editor, *INTETAIN 2005*, LNAI, pages 122–131. Springer, 2005.

[8] A. E. Scheflen. Micro-territories in human interaction. In A. Kendon, R. M. Harris, and M. R. Key, editors, *World Anthropology: Organization of Behavior in Face-to-Face Interaction*, pages 159–173. Mouton, Paris, 1975.

[9] W. Shao and D. Terzopoulos. Autonomous pedestrians. In *SCA '05: Proceedings of the 2005 ACM SIGGRAPH/Eurographics symposium on Computer animation*, pages 19–28, New York, NY, USA, 2005. ACM Press.

[10] H. H. Vilhjalmsson and J. Cassell. Bodychat: autonomous communicative behaviors in avatars. In *AGENTS '98: Proceedings of the second international conference on Autonomous agents*, pages 269–276, New York, NY, USA, 1998. ACM Press.