Natural Negotiation for Believable Agents

W. Scott Reilly Joseph Bates

June 1995
CMU-CS-95-164

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

This work was supported in part by Fujitsu Laboratories, Ltd. and Mitsubishi Electric Research Labs. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of Fujitsu Laboratories, Mitsubishi Electric Research Labs or any other parties.

Keywords: artificial intelligence, art and entertainment, believable agents, believable social behavior, interactive media, Oz

Abstract

Believable agents will often need to engage in social behaviors with other agents and with a user. Believable social behaviors need to meet a number of requirements: they must be robust, they must reflect and affect the emotional state of the agent, they must take into account the interpersonal relationships with the other behavior participants, and, most importantly, they must show off the artistically defined personality of the agent. We will show how to create a negotiation behavior that supports believability for specific characters and address some methodological questions about how to build believable social behaviors in general.

1 Introduction

The Oz project at CMU is developing tools that will allow artists to create dramatically interesting interactive media. As in traditional artistic media (such as literature, film, and character animation), an important part of such systems will be the characters (agents). Essentially, we want the rich characters of traditional artistic media combined with interactivity. We call these interactive characters *believable agents.*

Over the past few years "believable agents" research has emerged in the AI, arts, and entertainment community[1]. The goal of this work is to create autonomous interactive agents that exhibit qualities associated with believable characters in the arts. Many of the techniques we use come from artificial intelligence (AI), but AI has traditionally focused more on rationality and optimality in problem solving than on things like personality and emotional behavior. In fact, many excellent characters in traditional media are far from optimal problem solvers, so the focus of traditional AI needs to switch to be applied to this new domain.

Also, we are not necessarily interested in lifelike or realistic agents. We don't really care if our characters only have four fingers or are talking animals. The arts have always abstracted away from reality and created characters that were "larger than life" and that were much more interesting than any mere "real" people or animals we might know. By "believable agents" we mean interactive versions of these abstracted characters that have been so successful in non-interactive media.

The qualities of traditional characters that produce a suspension of disbelief in viewers are understood informally by artists (see, for instance, [11]). To produce believable autonomous agents, however, these notions must be expressed as computer programs. Believable agents researchers are working toward this computational modeling of artistic content. The work is proceeding at varied AI research centers, including MIT, Stanford, Carnegie Mellon, and numerous corporate laboratories.

Most believable agents will have to interact with other agents and human users. The work reported here is an attempt to create agents that negotiate believably. We call this "natural negotiation" to avoid confusion with work

[1]See especially the AAAI 1994 Spring Symposium on Believable Agents, and the believable agents session in the AAAI-94 national conference.

on negotiation in Distributed AI (DAI) such as [10]. Studying natural negotiation both provides insights into an important social behavior and allows us to investigate possibly useful methodologies for creating other believable social behaviors, such as threatening, order giving and following, and initiating relationships.

As in our previously reported work [1, 6], emphasizing believability leads us in new directions. The arts suggest that the most crucial element for believability is to imbue characters with strong personalities that permeate their action. In contrast, rationality and optimality, goals of much DAI negotiation research, are not only secondary, but would likely break the believability of many artistically effective agents. The need to have our agents express personality will surface repeatedly in the ideas presented here.

In addition to personality, there are several other aspects of characters important for believable social behavior. First, agents need to appear to have emotional reactions to their experiences. Second, behaviors need to be robust enough to not fail in situations that can reasonably happen within their environment. Third, agents need to be able to have dynamic social relationships that affect the agent's behavior. Fourth, agents need to appear competent at working towards multiple goals simultaneously.

After presenting some of our underlying methodological and technical assumptions, we will look at these aspects of believability as they relate to negotiation.

2 Foundational Work and Methodology

2.1 Foundations

The work being described here has been done in the context of the Oz system. The Oz project's goal is to enable artists to create dramatic microworlds that contain believable agents with which human users can interact. We hope this form of interactive art has the potential to provide users with deep dramatic experiences.

We have previously argued [1] that an effective approach to building believable agents is to create broad agents. A broad agent includes a wide range of tightly integrated capabilities, such as reactivity, goal-directed behavior, emotion, memory, and language. We hypothesize that building any capabil-

ity to a deep level of competence will not be as important as creating a broad set of integrated capabilities.

Loyall and Bates developed the Hap language [6] as a foundation for broad agents. Hap is a behavior-based language in the spirit of Firby's RAP system [4] and Georgeff and Lansky's PRS system [5]. Like other behavior-based languages, Hap makes it natural to build behaviors that are robust and reactive. Unlike some other behavior-based frameworks, however, Hap uses explicit goals to organize behavior, which we feel helps give our agents the appearance of intention.

We have previously built a number of simple, but perhaps successful, believable agents, including Lyotard the cat [1] and the Woggles [6]. These agents display robust, goal-directed behavior, they are reactive to their environment, and, unlike traditional AI agents, they act emotionally and with distinct personalities. Informal evidence from users interacting with these agents (hundreds of users in the case of the Woggles) suggests that we are progressing towards believability.

2.2 A Methodology for Building Social Behaviors

Even though we knew we would be using the Hap language for creating social behaviors, there were still unanswered questions about how to proceed. The biggest question was: what kinds of representations do agents need to have of other agents in order to engage in social behaviors?

Instead of trying to answer the question outright, we adopted what we believe to be a reasonable methodology. We are building behaviors using the minimal amount of representation sufficient to create those behaviors. This approach is inspired by the work of Brooks [2] who showed that surprisingly complex physical behaviors could be created with little or no representation. He also argued that for complex, dynamic domains, maintaining accurate internal representations is very difficult and will often lead to inappropriate behavior if relied on too heavily.

Mataric [7] extended Brooks's work to simple social behaviors for robots, such as flocking. We are following in this tradition but extending the types of behaviors and emphasizing the need for the behaviors to be believable as well as robust.

Others (such as Cesta and Miceli [3]) are using more complex knowledge representations, such as modal logics, for reasoning about other agents in

social domains. One of our goals is to explore the benefits and drawbacks of our minimalist methodology.

Another related question is: how do we represent an agent's relationship to other agents? For instance, how do we encode that Bob is friends with Sue? Because we wanted changing relationships, we felt that using explicit representations would be better than forcing all of the necessary information implicitly into behaviors. We also wanted to keep the representation as simple as possible. We settled on structures that encode the type of the relationship (e.g., friends) and the level of that relationship (from 0 [not friends] to 10 [very close friends]). Bob and Sue would each have such a representation, so Bob could think that he and Sue were better friends than she feels they are. This trivial encoding has, so far, proved sufficient for our needs.

3 Natural Negotiation

In this section, we describe what we mean by natural negotiation and how to create believable agents that engage in such behavior. We begin generally, with a description of what we mean by natural negotiation. Then we show two believable agents that engage in natural negotiation and describe how we created these behaviors. In section 3.4, we examine the representations agents might use of other agents in natural negotiation.

3.1 The General Behavior

Natural negotiation is simply negotiation for believable agents. This means, however, that the behavior must take into account the personality of the agent, the emotional state of the agent, and the interpersonal relationships the agent has with the other negotiator(s).

Two other requirements for natural negotiation that are not unique to believable agents are that it be robust and that it take the agent's other goals into account.

Robustness does not mean quite the same thing for believable agents that it does for other kinds of agents such as robots. Robots need to be deeply competent within their domain of expertise, but failure is acceptable for problems outside their area of competence. There can also be a hard edge between what a robot can and cannot do. A robot designed to explore Mars

doesn't need to know what a newspaper is since this is outside its range of expertise. If we ask a Mars rover what a newspaper is, we don't expect a reasonable reply.

Believable agents are simpler than robots because they generally don't have to be deeply competent in any area and because users typically come to our agents ready to suspend their disbelief. However, robustness is also harder for believable agents because the transition between what the agent can and cannot do must be gradual. For instance, if a user asks a believable agent what a newspaper is, they either need to know or know how to avoid the question believably. For instance, the agent could say, "I'm really busy, can we talk later?" or "Get out of my face, dork!" depending on its personality, emotional state, current goals, and relationship to the asker.

Believable behaviors also have to take other goals into account. In the simulated world that we describe in section 3.2, one agent is negotiating with the user and trying to become the user's friend. This second goal should affect the negotiation. Solutions to this problem are still speculative, so we leave further discussion to a future paper.

Each negotiating agent needs a behavior designed for that agent's personality, the world the agent will inhabit, and the kinds of negotiations the agent needs to take part in. It is unclear how to solve this problem in a general way, so our solution is to analyze the general behavior well enough that specific instances can be built with greater ease. We begin this analysis by noting three dimensions of complexity of the negotiation behavior.

Number of agents involved. One dimension of complexity is the number of agents involved. The simplest is two and it can grow without bound. Some multi-agent negotiations can be created out of two-agent negotiations, but not all.[2] One step up in complexity is to put together links of two agent negotiations. Another step is to be able to carry on true multi-agent negotiations for either set numbers of agents or arbitrarily many agents.

Object of negotiation. Another dimension of complexity is the type of thing being negotiated for. In our simulated worlds where sensing is easy and reliable, negotiating for visible objects is simpler than having to model the other agent's possessions or asking about items the other agent is willing to trade. Other types of things that can be negotiated for include: services,

[2]A three-way circular trade requires the agent to negotiate with two other agent simultaneously.

knowledge, and relationships.

Types of offers made and understood. A third dimension is the types of offers that the agent can make and understand. For example, an agent might be able to respond to offers from other agents but not be able to make offers. Another source of variation is whether the offers offered/understood include partial offers (e.g., "What will you give me for X?" or "What do you want for Y?") or full offers (e.g., "Would you give me Y for X?").

When designing a negotiation behavior for a specific agent, our first step is to decide where along these dimensions the behavior should fall. Interactions with agents that have complicated behaviors can be more interesting, but there is no need to create more complexity than is necessary for the environment in question and the personality of the agent.

Now we will look at a clean behavior for a simple kind of negotiation. By "clean" we mean that the behavior is technically correct, but the aspects we feel are most important to believability, such as personality and robustness, have not yet been added. We will explain how such a simple behavior is modified to add in the other aspects later. This behavior in Figure 1 handles negotiations for most kinds of objects with one other agent. Behavior A is for initiating the negotiation. Behavior B is for responding to an offer. Similar behaviors can be made for initiating and responding to partial offers as well.

3.2 Negotiating Agents: Melvin and Sluggo

Having sketched a space of negotiation behaviors, we will show how this kind of analysis is useful for creating specific negotiating agents. First, we describe refinements of the general behavior for two agents. In the next section we show how these refinements were made to produce the various aspects of believability.

The world we built is a simulated playground where the user acts as a child trading baseball cards with other children. The other children are Melvin, a friendly nerd who is competent at negotiation but who is as interested in the social interaction as with the outcome of the negotiation, and Sluggo, a bully whose lack of intelligence is made up for by his belligerence.

As described above, the first step in creating our agents was to define what kinds of negotiation they needed. Along the dimension of who to negotiate with, we decided that interactions with single other agents were fine. More complex negotiations might make them seem too mature and might be more

Behavior A: Agent initiates a full negotiation

```
1. Choose negotiation partner
2. Make offer
3. If offer accepted then make trade
   else if offer rejected without counter offer
        then fail or backtrack or persuade partner to accept
   else if offer rejected with counter offer
        then goto B.1
```

Behavior B. Other agent initiates a full negotiation

```
1. Evaluate offer
2. If offer acceptable
        then accept offer and make trade
   else if offer unacceptable and reasonable
        then make counter offer and goto A.3
   else if offer unacceptable and unreasonable
        then fail or backtrack
```

Figure 1: Clean Negotiation Behaviors

confusing for the user than interesting. Along the dimension of objects to negotiate for, the obvious choice was (visible) baseball cards. Finally, we decided that Melvin should be able to offer and respond to single and multi-card full (e.g., "I'll give you Ruth for Mays.") and partial (e.g., "What do you want for Aaron and Jackson?") offers, but that Sluggo's intelligence should only allow him to respond to one-for-one full offers or single card partial offers. Note that Sluggo's behavior is somewhat simpler than Melvin's. The complexity of the behavior is overshadowed by the importance of getting across Sluggo's personality, which comes across better through a simpler behavior.

We also built behaviors to handle the swapping of cards once the negotiation has ended in a mutually acceptable proposal. This part of the behavior also needs to reflect the personalities of the agents. We will show how this was done below.

Figures 2, 3, and 4 show a simple, annotated trace from a user's interaction with Melvin and Sluggo.[3] Since this comes from an interactive system, this is only one of many possible courses such a trace could take. It is meant to give the reader a feel for the characters we have built and to show a bit of their personalities, emotions, interpersonal relationships, and robustness. The next section explains how many of the aspects of the behavior described in the trace were accomplished.

The user is acting by typing at the "PLAYER>" prompt. Information not at the prompt is system–generated descriptions of what the user can see and hear. The trace has been lightly edited and annotated for clarity.[4]

3.3 Believability Issues in Negotiation

In this section we will look at a number of important aspects of negotiation necessary for believability, including robustness, emotion, personality, and interpersonal relationships. For each, we will discuss some heuristics for adding it to a "clean" behavior like the one introduced in section 3.1 to produce behavior like that of Melvin and Sluggo. We are not trying to provide a comprehensive analysis of the ways that these aspects of believability interact with natural negotiation. Instead, we want to show that there are many such interactions, they appear at all levels of the behavior (from top-level goals to plans to physical actions), and they can be achieved using proper organization, artistic techniques, and the simple mechanisms provided in a language like Hap.

Robustness

Robustness is how well a behavior stands up to the unexpected. For instance, Melvin and Sluggo have to be ready for the user to wander away at any time or make offers unrelated to what had gone before. None of these are supported in a behavior like that in Figure 1. The Hap language provides support for robustness in social behaviors the same way it does for other

[3]This simulation was built in the Lisp version of the Oz system. The system has roughly 50K lines of code and each of the agents described here accounts for roughly 3K. Running on an HP-9000/735, each agent's action selection takes roughly 1-2 seconds.

[4]We make no claims about the language capabilities for the agents. The language generation uses templates and understanding uses pattern matching of keywords.

```
You are in the playground.

   The sand box, the jungle gym and the tree house are in the playground.

   Sluggo is in the tree house.
   Sluggo is holding a Willie Mays trading card, a Jose Canseco
   trading card and a Catfish Hunter trading card.

   Melvin is in the sand box.
   Melvin is holding a Tom Seaver trading card, a Mickey Mantle
   trading card and a Reggie Jackson trading card.
   Melvin is wearing his eye glasses.

   You are holding a Babe Ruth trading card, a Ted Williams
   trading card and a Henry Aaron trading card.

PLAYER> get in the sandbox

   Melvin is now smiling.
   Melvin says to you: ''Greetings, earthling.''.

PLAYER> Say to Melvin: Wanna trade?

   Melvin says to you: ''The Klingon Emporer wants to know if you
        would be willing to part with Babe Ruth for Reggie Jackson?''.
```

[Melvin wants to become the Player's friend. One way to accomplish this is to interact socially through negotiation. In this case, the interaction is more important than the outcome of the negotiation. Melvin starts off happy because his goals of being near the player and interacting with the player are being met. Also, we can already see some of Melvin's unique personality starting to show through.]

```
PLAYER> Say to Melvin: No thanks.

   Melvin says to you: ''Captain Picard would make the trade.''.
```

[Melvin's offer is refused, so he tries to be persuasive.]

```
PLAYER> Say to Melvin: Sorry, no deal.

   Melvin says to you: ''Are there any cards we (the Klingon high
        command and I) can offer you in return for Babe Ruth?''.

PLAYER> Say to Melvin: Do you have anything that would help me get Mays from Sluggo?

   Melvin says to you: ''So the Vulcan ambassador wishes to get Willie
        Mays from Darth Sluggo.  He really likes Mickey Mantle you
        know and I'd trade you.''.
```

[The Player rather abruptly changes the course of the negotiation by asking for help getting a card from Sluggo. Melvin is able to react appropriately. In this case, Melvin's goal to become friends with the player causes Melvin to help and he does so by making an offer in both Melvin's and the Player's interest.]

Figure 2: Trace of Melvin and Sluggo Negotiating (Part 1)

```
PLAYER> Say to Melvin: Okay, how about Mantle for Ruth?

   Melvin says to you: ''The Klingon high command accepts Vulcan's generous offer.''.
```

[A deal is reached and the swapping of cards begins. Because Melvin wants to become friends with the Player, Melvin would accept any reasonable offer, even ones that Melvin didn't consider a "good" deal. In this case, however, he gets a card that he wanted from the beginning.]

```
PLAYER> Give Ruth to Melvin

   Melvin takes a Babe Ruth trading card.

PLAYER> wait

   Melvin offers a Mickey Mantle trading card to you.

PLAYER> take the card

   Melvin says to you: ''It's a pleasure to hang out with you, earthling.''.

PLAYER> Say to Melvin: I'm going to go talk to Sluggo.  Bye.

   Melvin says to you: ''You are welcome on Klingon anytime!''.
```

[Melvin is now quite happy because of the successful interaction and the fact that he got a good trade. He compliments the player and then they part amicably.]

```
PLAYER> get in the tree house

   Sluggo says to you: ''Bud, wanna trade Mickey Mantle for Catfish Hunter?''.
```

[Sluggo isn't particularly interested in a social interaction. He just wants some good cards.]

```
PLAYER> Say to Sluggo: How about Mantle and Aaron for Hunter and Canseco?

   Sluggo is now red.
   Sluggo is now scowling.
   Sluggo is now tense.
   Sluggo says to you: ''You think you're cool with your fancy trades?  No deal, dweeb!''.
```

[Sluggo can't handle such a complex offer, but he doesn't want to admit it, so he rejects what would have been a good trade for him. Sluggo gets a bit angry that the player made an offer that he couldn't understand, but not enough to really affect the behavior. If the Player continued to make such offers, Sluggo would start to show his anger.

Even when Sluggo understands an offer, he will only accept it if it is clearly in his best interests.]

Figure 3: Trace of Melvin and Sluggo Negotiating (Part 2)

```
PLAYER> Say to Sluggo: What'll you give me for Mickey Mantle?

   Sluggo says to you: ''I'll swap ya Jose Canseco for Mickey Mantle.''.

PLAYER> Say to Sluggo: No way!

   Sluggo says menacingly to you: ''I think you should reconsider, bub.''.
```

[Sluggo can also be persuasive in his own way.]

```
PLAYER> Say to Sluggo: Um, OK.  I guess I'll make that trade after all.

   Sluggo offers a Jose Canseco trading card to you.

PLAYER> Take card

   Sluggo says to you: ''So, hand over the card, twerp.''.
```

[Sluggo is rather impatient and demanding.]

```
PLAYER> Say to Sluggo: Hey, lighten up!

   Sluggo looks angrily at you.

PLAYER> Look angrily at Sluggo

   Sluggo says to you: ''You don't know who you're messin' with, dork!''.

PLAYER> Say to Sluggo: You don't scare me.

   Sluggo is now frowning.
   Sluggo says to you: ''Prepare to be pounded!''.
```

[As the player continues to be uncooperative, Sluggo gets more angry, until he resorts to threats and then violence. Fighting, however, is a social behavior for another time...]

Figure 4: Trace of Melvin and Sluggo Negotiating (Part 3)

kinds of behaviors, by means of preconditions, demons with varying priorities, success-tests, and context-conditions, which are described elsewhere [6]. Using these mechanisms, we need to organize the behaviors to provide robustness. There are two important ways to organize behaviors to be robust: be complete and be ready when you aren't complete.

Being Complete. One way to be robust is to build the behavior with as many contingencies as possible. When waiting for another agent to respond to an offer (Figure 1, A.3), the behavior must be ready for as many kinds of responses as possible. The responses Melvin and Sluggo recognize are: yes, no, a counter-offer, handing over a card in the offer (assumes a "yes"), a new full offer, both kinds of partial offers, and offers that cannot be made based on who currently owns what cards are also supported. Melvin and Sluggo can also move between behaviors when necessary, as demonstrated when Melvin's offer to the player is followed by a request from the player for help, which triggers a whole new behavior in Melvin. By including as many responses as possible and the ability to interrupt behaviors with other behaviors, we make social behavior more robust.

Being Ready for the Unexpected. Another way to provide robustness is to provide general responses for situations that fall outside of the domain of the behavior. Our agents cannot possibly be ready for every situation, but they can't seem confused at inappropriate times either. The clean behavior does not support this kind of robustness. When Melvin makes an offer and the user does something unexpected, Melvin will wait or maybe whistle. Maybe the user is looking over Melvin's cards or maybe she just mistyped a command. In these cases, Melvin's generic actions may be reasonable and allow the behavior to get back on track when the user is ready. If the user has ended the negotiation and begun a new behavior, Melvin will whistle briefly then begin a new behavior. This means Melvin doesn't get stuck waiting on a response that never comes, but neither is the behavior so brittle it falls apart when small interruptions occur.

A similar kind of situation arises in language processing. Melvin and Sluggo are programmed to respond to a few statements and requests, but the user could potentially talk about anything. In these situations, we need to have a reasonable, but generic response. For example, here's one response we can get from Sluggo:

```
PLAYER> Say to Sluggo: How about them Steelers?
   Sluggo says to you: ''Hey dork, shut up until I tell you to talk.''.
```

In this case, Sluggo has no idea what the player has said, but instead of breaking the suspension of disbelief, we are able to turn Sluggo's response into an outlet for his personality.

Emotion

Building emotion into behaviors is simplified by the fact that we are using an action architecture that is already coupled to an emotion module (Em) [9]. If things are set up properly, Em automatically generates emotions and helps express them, which takes that burden off of the agent builder. Nonetheless, there is still work to do.

The coordination of emotion and behavior has two components. First, the behavior may be the source of emotions. We need to put information in the behavior to let Em generate such emotions. Second, emotions may affect behavior. This is the more difficult of the two for the behavior builder. We will look at each of these.

Generating Emotions from Natural Negotiation. The Em system uses a set of default emotion generation rules based on the emotion model of Ortony, Clore, and Collins [8]. To generate emotions, these rules look at (among other things) simple annotations added to important goals that include: the importance of success, the importance of not failing, a likelihood of success function, a likelihood of failure function, a function to determine who is responsible if the goal succeeds, and a function to determine who is responsible if the goal fails.

In the trace, the user stalls while Sluggo is expecting his cards. Sluggo's Em module finds his important `get-cards`[5] goal is more likely to fail, which leads to fear, distress, and anger towards the player at an intensity based on the importance of the goal not failing and the likelihood that it will fail. Because Sluggo also believes that this goal may still succeed, Em also generates hope. Which emotion most affects behavior is based on their relative intensities and the personality of the agent. Sluggo will tend to show negative emotions, like fear and anger, more that positive emotions, like hope. He also tends to show negative emotions through aggression. Em also generates joy and gratitude when the trade is successful or distress and anger when the other agent fails to fulfill the bargain.

[5] A subgoal of the `make-trade` goal in Figure 1 A.3 and B.2.

The Effect of Emotions on Natural Negotiation. The negotiation behavior needs to reflect the agent's emotional state, whether that state is a result of negotiation-related emotions or not. There are both general and negotiation-specific ways that emotions influence negotiation. Three general effects are changes to appraisals, changes to goal priorities, and changes in willingness to interact. These types of changes are general in that they can be seen in other social behaviors as well.

1. Emotions can changes how an agent appraises objects (see Figure 1, B.1). Sad agents may be willing to make trades they otherwise wouldn't have made because they may appraise the value of their possessions less highly. There are a few types of appraisals in our agents, but they are mostly generic functions that can be written to take emotion information into account. Remember that rules of thumb, like "sad agents appraise their possessions less highly" are in no way enforced by the system. In fact, it is important that artists be able to create characters that break any such general rule about typical behavior. We don't want typical agents, we want interesting, unique agents.

2. An example of a change in goal priority comes about when an agent is afraid that the goal to get the cards from a trade is in jeopardy. This is normally a somewhat high priority goal, but if the agent is afraid that this goal will fail, the agent will probably be very persistent about achieving this goal before doing anything else. Fear of other goals will lead to responses appropriate for those goals, such as running away in response to some health-preservation goals. To accomplish this, we extended the Hap language to allow dynamic priorities for goals, so a goal's priority can be a function that takes relevant emotion information into account.

3. Emotions influence how likely an agent is to interact. Happy agents, for example, are usually more likely to interact than usual; sad agents are less likely to interact. Emotions affect how likely the agent is to initiate a negotiation and how the agent responds to offers from others. These effects are expressed in the preconditions for various plans. Melvin has a `respond-to-offer` demon that fires when another agent makes an offer and kicks off the appropriate behavior. Behaviors that are more socially responsive have preconditions that make sure that Melvin is in the proper emotional state before firing.

There are also ways that emotions influence negotiation specifically. For instance, an agent feeling fear may be less likely to engage in negotiations for

important objects. Agents that are feeling gratitude towards another agent will likely be more generous toward that agent; agents that are angry at an agent will be less generous. All of these effects can be produced through methods similar to the ones described in the three examples above or through the normal hap constructs, like demons with varying priorities, preconditions, context-conditions, and success-tests. For example, changes in generosity can be coded either in the appraisal function or in the precondition for the `accept-offer` and `make-offer` behaviors.

Agent builders need to fit these emotion–negotiation interactions to their specific agents. We have sketched some heuristics, and not a complete set at that.

Interpersonal Relationships

The relationship between two agents should affect the way the negotiation unfolds and, when things go particularly well or poorly, the relationship itself should change based on the negotiation. Also, agent's attitudes about other agents influence and are influenced by negotiation. To clear up any confusion about terminology, relationships are things like "friends" and "lovers", while attitudes are things like "trust" and "liking". While these are different concepts, the way they relate to negotiation is similar.

The Influence of Relationships and Attitudes on Negotiation. As the following examples suggest, relationships and attitudes influence negotiation at all levels of the behavior, from the highest level goals to the choices of specific actions:

1. The precondition for Melvin's `make-offer` behavior (Figure 1, A.2) prefers initiating negotiations with friends and liked agents and filters out negotiating with enemies and feared or disliked agents.

2. When deciding how to respond to an offer (Figure 1, Behavior B), Melvin's `respond-to-offer` and `respond-to-partial-offer` demons reject offers from enemies and feared or disliked agents.

3. When speaking to another agent, the tone can change while the message remains the same. When Melvin accepts an offer, he can say, "Okay." or "Yeah, that sounds good." or "Sure! I'd love to make that trade!" Each of these has the same basic content, but which is chosen is still very important and the choice is influenced by the relationship between Melvin and the other agent.

The Influence of Negotiation on Relationships and Attitudes. If Melvin considers the user his friend, but the user cheats him, Melvin should probably no longer consider himself and the user friends. Their new relationship affects future negotiations as described above. This also holds for attitudes. Melvin, in fact, does not consider the player his friend, but he does like the player and want to become friends. If the player cheats him, Melvin will like the player less, which, in turn, affects furture interactions.

In addition to Melvin and Sluggo, we have also built a number of other agents that adjust their attitudes and relationships over time, such as Lyotard the cat [1] and the Woggles [6]. We have found that the simple type-value representation of relationships described in section 2.2 has been useful for creating dynamic relationships.

Personality

Personality is the most important aspect of believability to add to social behaviors. We have found that personality permeates behaviors; almost every element of a behavior can be an opportunity to represent personality. This is one reason we have focused on techniques for building specific behaviors for particular agents instead of trying to build a generic negotiation behavior.

Each of the previous three sections (robustness, emotion, and interpersonal relationships) can be considered important parts of the way to get personality into negotiation. They don't, however, cover all of the ways. Here are a few additional ways we expressed Melvin and Sluggo's personalities in their negotiation behaviors:

1. *Handling uncooperative trading.* Melvin and Sluggo approach uncooperative trading parters differently. Melvin is persistent but polite: if pushed too far, he will eventually give up but hold it against the player. If Melvin is thwarted again, he will tell the teacher and get the player in trouble. Sluggo goes into the trade with an impatient and aggressive attitude and quickly moves towards more aggression and threats which finally end in violence. This is best achieved by writing two distinct behaviors for handling uncooperative agents in Melvin and Sluggo.

2. *Understanding offers.* We decided that Sluggo should be rather dumb. One reason for this was to show that our goals are different from those of traditional AI where intelligence and rationality are important. One way we made Sluggo less intelligent was to make him unable to understand complex

offers in a negotiation. This was done by having the precondition for accepting offers filter out offers that are too complex. Melvin has no problem with this task.

3. *Appraisals.* Different characters should like and want different things. This is reflected in their appraisal functions. Sluggo really likes Jackson and Mantle. Melvin likes Babe Ruth.

4. *Generic actions.* There are a number of situations where Melvin or Sluggo need to kill a turn or two when nothing else is going on or when they are waiting on another agent. Melvin will whistle, mumble, or look intently at some cards. Sluggo uses this time to spit, swear, and look intently at the other agent. Even these little, seemingly unimportant actions have a good deal of impact on the way the user views an agent and their importance in creating believability should not be overlooked. Most of these actions were edited from the trace for space reasons.

3.4 Representation Issues in Natural Negotiation

In addition to building a useful behavior for believable agents, we are also trying to shed light on the question of how much representation of other agents is sufficient to build such a behavior. For Melvin and Sluggo, the only representation used were lists of the offers other agents had rejected in the past so as not to repeat them. With this minimal amount of representation, we believe the behavior is robust enough and the agents display enough competence to be believable in this domain.

This doesn't mean that other representations won't ever be necessary to create competent negotiation. In fact, we can imagine at least two types of representation that would be useful in some domains and for some personalities: a representation of trustworthiness of the other agent and a representation of the appraisal function of the other agent.

Trust is useful in domains where agents deceive each other. The playground agents don't run into this problem because if the user goes back on a deal they deal with it in ways that preclude further negotiations (either getting the user in trouble with the teacher or beating up the user).[6] If there

[6]Actually, Melvin can be tricked once, which leads to him liking the player less, but which does not affect how likely he is to trade with the player a second time. This makes him seems a bit gullible and over-trusting, which seemed to fit his personality fine. If he is cheated a second time, he will tell the teacher.

were a possibility for future interactions with such a deceitful user (or other agent), decisions about who to deal with and how to appraise offers from deceitful agents need to take trust into account. We hypothesize that trust can actually be represented adequately by a single integer value. When an agent fails to follow through on a promise, this number should drop to a low value. Positive trade outcomes should raise such a number, but gradually. We believe that such trivial representations are often sufficient for believability.

A second useful type of representation would be a model of the appraisal function of the other agent. If Melvin has an object that he doesn't consider very valuable, he might trade it for very little. However, if he knows that Sluggo really wants the object, he can use that knowledge to get more out of the trade (depending, of course, on his personality, emotional state, and relationship with Sluggo). This requires Melvin to have a model of Sluggo's appraisal function that is different from his own.[7] A model of an appraisal function that provided this level of competence could be quite simple in most cases, such as a list of object-value pairs.

Another level of complexity comes from basing offers on the goals of the other agent. If the user were dealing with a villain that was trying to hinder the user, that agent might negotiate with the user, but certainly not want to give the user anything necessary to achieve any important goals. This means that the model of the user's appraisal function might have to take the user's dynamically changing goals into account.

In some cases, representations of other agents' goals can be avoided. If the villain were pretending to help the player, he could ask the player about his/her goals and make the appraisal accordingly. In other environments, the player's goals might be known by the agent builder because of the specifics of the environment and (in story-based systems) the plot, so the appraisal function model could be written with this knowledge contained implicitly. These methods do not cover all cases, however, and sometimes models of other agents' goals will be needed. Even in this case, however, it should be possible to use simple goal representations that do not require deep modeling of the goal processing of the other agents. A list of goal tokens (e.g., get-

[7]Of course, Melvin could ask Sluggo about his appraisal of the object, which would mean that Melvin didn't need to model Sluggo's appraisal function. This approach would lead to unbelievable behavior, however, in cases where Melvin should know this information and Sluggo could possibly lie.

treasure) or goal lists (e.g, (rescue bob :importance 6)) will often be enough to create the desired behavior. We believe that using internal simulations of other agents' goal processing will be (at least in most cases) unnecessary.

We feel our work with natural negotiation and other social behaviors in a variety of simulated worlds has shown our minimalist methodology to be effective. We will continue to explore the usefulness of this methodology.

4 Discussion and Future Work

The work presented here is one step towards the goal of creating believable agents that engage in social behaviors. Creating a natural negotiation behavior has not only been an end to itself, but a way to test our methodology and underlying assumptions. In this section, we will discuss how each of the following choices influenced our ability to build a believable social behavior: (1) using a behavior-based architecture, (2) using minimal representations of other agents, and (3) focusing on personality, emotion, robustness, and relationships.

The choice to use the behavior-based Hap language proved to be important for providing robust, reactive behavior. Hap was primarily designed to be a language for creating reactive physical behaviors, but the same kinds of structures that are necessary for robust physical actions also turn out to be useful for social behaviors. Hap needed to be extended to give us some new kinds of reactivity, but the basic behavior-based structure has served us well.

Using minimal representations has also been surprisingly effective. Brooks was able to eliminate representations in his robots by using the physical world as its own model. There is, however, no way to sense the internal state of other agents, so it wasn't obvious Brooks's approach would work in the social world. We feel we were able to build robust negotiation with the only representation being what offers the other agents had rejected. As discussed above, more representation might be needed in other situations, but there does not yet seem to be a need to move towards a more powerful representation language such as modal logic.

The aspects of believability we have focused on have come mostly from artists in other media. Personality, emotion, and interpersonal relationships between characters are important to characters in novels, movies, and plays so it isn't surprising that they are also important in interactive artistic me-

dia. Robustness, however, is something that isn't applicable to characters in traditional media. We have found that agents that are unresponsive to the user are not only unbelievable, but frustrating to interact with. ELIZA [12] is an example of a robust character that seems believable to many people, but doesn't show the personality that makes artistic characters memorable. We want to add enough robustness that people can interact naturally, but not at the cost of other aspects of believability.

Although we had some difficulties in creating natural negotiation, none appears to have been based on fundamental flaws with our underlying assumptions or methodologies. We believe our approach will apply as we create other believable social behaviors.

Two near-term goals for this work are to develop a way to evaluate our work in an effective manner and to build more believable social behaviors. Evaluating our work is an interesting issue. Because our goal is that of believability, traditional evaluation metrics of speed and correctness do not seem to apply. We expect to use some sort of user testing, but the details of such testing remain unsettled. Part of the difficulty with evaluating work in this field is that we need to take both the technical and artistic elements into account. We expect that as the research area of believable agents expands, methods for evaluating work within the field will also develop.

References

[1] Joseph Bates, A. Bryan Loyall, and W. Scott Reilly. An architecture for action, emotion, and social behavior. In *Proceedings of the Fourth European Workshop on Modeling Autonomous Agents in a Multi-Agent World*, S.Martino al Cimino, Italy, July 1992.

[2] Rodney Brooks. Intelligence without representation. In *Proceedings of the Workshop on the Foundations of Artificial Intelligence*, June 1987.

[3] Amedeo Cesta and Maria Miceli. In search of help. In *Proceedings of the 12th International Workshop on Distributed Artificial Intelligence*, Hidden Valley, PA, May 1993.

[4] James R. Firby. *Adaptive Execution in Complex Dynamic Worlds.* PhD thesis, Department of Computer Science, Yale University, 1989.

[5] Michael P. Georgeff, Amy L. Lansky, and Marcel J. Schoppers. Reasoning and planning in dynamic domains: An experiment with a mobile robot. Technical Report 380, Artificial Intelligence Center, SRI International, Menlo Park, CA, 1987.

[6] A. Bryan Loyall and Joseph Bates. Real-time control of animated broad agents. In *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society*, Boulder, CO, June 1993.

[7] Maja J. Mataric. Kin recognition, similarity, and group behavior. In *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society*, Boulder, CO, June 1993.

[8] A. Ortony, G. Clore, and A. Collins. *The Cognitive Structure of Emotions*. Cambridge University Press, 1988.

[9] W. Scott Reilly and Joseph Bates. Building emotional agents. Technical Report CMU-CS-92-143, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, May 1992.

[10] Katia Sycara. Resolving goal conflicts via negotiation. In *Proceedings of the Sixth National Conference on Artificial Intelligence*, pages 245–250, St. Paul, MN, July 1988.

[11] Frank Thomas and Ollie Johnston. *Disney Animation: The Illusion of Life*. Abbeville Press, New York, 1981.

[12] Joseph Weizenbaum. Eliza. *Communications of the ACM*, 9:36–45, 1966.