Dealing with Doctors: A Virtual Human for Non-team Interaction

David Traum and **William Swartout** and **Jonathan Gratch** and **Stacy Marsella**
Patrick Kenny and Eduard Hovy and Shri Narayanan and Ed Fast
Bilyana Martinovski and Rahul Baghat and Susan Robinson and Andrew Marshall
Dagen Wang and Sudeep Gandhe and Anton Leuski
University of Southern California
Marina del Rey, CA, USA
traum@ict.usc.edu

Abstract

We present a virtual human doctor who can engage in multi-modal negotiation dialogue with people from other organizations. The doctor is part of the SASO-ST system, used for training for non-team interactions.

1 Introduction

Virtual humans can play an important role in helping train skills of interacting with others who have different beliefs, goals, and styles of behavior. By building virtual humans that are not just humanoid in appearance and external behavior, but which also have internal models (including beliefs, goals, plans, and emotions) and ability to reason over these models and formulate appropriate strategies and behaviors on the basis of the models and perceptual input, virtual humans can behave appropriately for a range of social relationships.

In previous work (Rickel et al., 2002; Traum et al., 2003), we described a negotiation model that could allow virtual humans to engage as teammates. This model assumed that teammates shared common end goals, participated in a social institution with roles that the participants played, and had strong trust in the other teammates' abilities and veracity. It did not address how virtual humans might interact in the case where these factors were lacking, and how to begin to form them through interaction.

In this demo, we present Dr Perez, a virtual human implemented with an extension to this model. The extended model allows for the case in which relationships may need to be developed during the interaction, and in which the virtual human's behavior may be very different depending on the nature and strength of the relationships. More details on the agent model can be found in (Traum et al., 2005b) and more on the negotiation model can be found in (Traum et al., 2005a).

Figure 1: VR clinic and virtual human doctor

In the next section, we describe our initial testbed: a scenario within the SASO-ST project. In Section 3, we briefly describe the virtual human model and how trust of the agent toward another is calculated. In section 4, we show two example interactions with this agent, showing how the dynamic trust model is developed during the interaction and how this can affect the agent's choice of utterance. We conclude with some remarks about current and future directions.

2 Domain Testbed: support operations

Whether it is Kosovo, East Timor, or Iraq, one lesson that has emerged from attempts at "peacemaking" is that negotiation skills are needed across all levels of civilian and government organizations involved. To have a lasting positive effect, interactions between military and locals must be carried out in a way that generates goodwill and trust. We have selected this general class of operations as a testbed for our work on negotiation.

More specifically, we are developing a training scenario in which a local military commander (who has

a rank of captain) must negotiate with a medical relief organization. A virtual human plays the role of a doctor running a clinic. A human trainee plays the role of the captain, and is supposed to negotiate with the doctor to get him to move the clinic, which could be damaged by a planned military operation. Ideally, the captain will convince the doctor without resorting to force or threats and without revealing information about the planned operation. Figure 1 shows the trainee's view of the doctor in his office inside the clinic. The success of the negotiation will depend on the trainee's ability to follow good negotiating techniques, when confronted with different types of behavior from the virtual doctor.

The success of a negotiation is also mediated by factors that influence the perceived trust between parties, including a belief in shared goals, credibility and interdependence. The doctor is unlikely to be swayed by an offer of aid if he does not believe the captain can and will fulfill his commitments. Trust issues are pervasive throughout the negotiation, since there is usually not much point in negotiating with someone you expect to lie, be ill-disposed toward you, or not keep their side of a bargain.

3 Virtual Human Negotiation Implementation

Our starting point was the virtual humans implemented as part of the MRE project (Rickel et al., 2002). These virtual humans are embedded in a dynamic virtual world, in which events can happen, agents can perform actions, and humans and virtual humans can speak to each other and communicate using verbal and non-verbal means. The virtual humans are extensions of the Steve agent (Rickel and Johnson, 1999), and include sophisticated models of emotion reasoning (Gratch and Marsella, 2004), dialogue reasoning (Traum and Rickel, 2002) and a model of team negotiation (Traum et al., 2003). Agents use a rich model of dialogue closely linked with a task model and emotional appraisals and coping strategies for both interpretation of utterances as well as for decisions about when the agent should speak and what to say.

To negotiate and collaborate with humans and artificial agents, virtual humans must understand not only the task under discussion but also the underlying motivations, beliefs and even emotions of other agents. The virtual human models build on the causal representations developed for decision-theoretic planning and augment them with methods that explicitly model commitments to beliefs and intentions. Plan representations provide a concise representation of the causal relationship between events and states, key for assessing the relevance of events to an agent's goals and for assessing causal attributions. Plan representations also lie at the heart of many reasoning techniques (e.g., planning, explanation, natural language processing) and facilitate their integration. The decision-theoretic concepts of utility and probability are key for modeling nondeterminism and for assessing the value of alternative negotiation choices. Explicit representations of intentions and beliefs are critical for negotiation and for assessing blame when negotiations fail (Mao and Gratch, 2004).

3.1 Modeling Trust

According to the dialogue model in (Matheson et al., 2000), the direct effect of an assertion is the introduction of a commitment, whether or not either party believes in the assertion. While this is sufficient for reasoning about the claims and responsibility for information, we need to go further and potentially change beliefs and intentions based on communicated information. Trust is used to decide whether to adopt a new belief based on the commitments of another.

Similar to (Marsella et al., 2004) and (Cassell and Bickmore, 2001) , trust is modeled as function of underlying variables that are easily derived from our task and dialogue representations. *Solidarity* is a measure of the extent to which parties have shared goals. It is derived from a running tally of how many times the trainee makes assertions or demands that are congruent with the agent's goals. *Credibility* is a measure of the extent a party makes believable claims. It is derived from a running tally of how many times the trainee makes assertions that are consistent with the agent's beliefs. Finally, *familiarity* is a measure of the extent to which a party obeys norms of politeness. Currently, an overall measure of trust is derived as a linear combination of these three factors.

3.2 Negotiation Strategies

In (Traum et al., 2005a) we describe the negotiation strategies that the virtual doctor uses, based on his current feeling about the desirability and avoidability of the object of negotiation, and the degree of closeness with his interlocutor. A strategy consists of several aspects including: **entry conditions**, which indicate when adoption is appropriate; **exit conditions**, which indicate when the strategy should be dropped (often in favor of more appropriate strategies); **associated moves**, which can be performed as tactics to implement the strategy; and **influ-**

ences of the strategy on behavior and reasoning. These aspects result from the underlying emotion and dialogue models of the virtual humans. Aside from rudimentary strategies for opening and closing a dialogue, we have implemented three negotiation strategies, patterned on studies of human negotiation (e.g., (Walton and Mckersie, 1965; Sillars et al., 1982)). If the doctor feels the negotiation is undesirable but avoidable, he chooses an *avoidance* strategy, manifested by attempts to leave the conversation or change the topic. If he feels it is undesirable yet unavoidable, he chooses an *attack* strategy, manifested by questioning the captain or pointing out (potential) problems in the plan. Finally, if the doctor thinks he can get something positive from the negotiation, he will adopt a **negotiate** strategy, in which case, operation is much like a team-negotiation (Traum et al., 2003).

4 Example Interactions

Consider the dialogue in Figure 2. This is just one of many possible interactions, depending on the choices of the human captain, as well as several aspects (some probabilistic) influencing the choice of moves and strategy transitions of the virtual human doctor.

Here the captain acts as he might with a team member - after starting the conversation, launching directly into his purpose, and answering a question straightforwardly. While this would have worked with a subordinate team-member, it has disastrous effects on the neutral doctor, bringing his trust level almost down to zero and failing to accomplish both objectives.

In this dialogue, nothing was done by the captain to try to establish a better relationship with the doctor, or address the issue of differing objectives and beliefs. The first exchange after the greetings (utterances 2-5) lowers solidarity by showing different objectives, setting up more of an antagonistic than cooperative interaction. The doctor tries to avoid the topic, focusing instead on his patients, rather than the captain's stated goal. The captain tries to argue for his proposed course of action, but only makes things worse with utterance 7. First, he says something the doctor doesn't believe (that the clinic is in danger), lowering his credibility. The doctor is able to reason though that perhaps the captain knows of a reason why it will be unsafe, and challenges by asking if he is going to cause the danger. In 9, the captain answers sincerely, which is a mistake on two fronts. First, he reveals more about his mission than he should to an outsider, possibly endangering it's success if word gets out to his enemies. Second, he shows even further divergence from the doctor's goals

1 C Hello Doctor Perez.
2 D Hello.
3.1 C I have orders to move this clinic to another location.
4 D You want to move the clinic?
5 C Yes
DECREASES SOLIDARITY: captain endorses undesired act 'run-clinic-there'
6.1 D Look at these people!
6.2 D we need to help them.
7 C It is not safe here, we can't protect you.
DECREASES CREDIBILITY: captain asserted unbelieved (but possible) state 'patients-unsafe-here'
'patients-unsafe-here' could be established by captain's act of 'planned-attack'
8.1 D Protect me? Protect me from what?
8.2 D Are you going to attack?
9 C Yes
DECREASES SOLIDARITY: captain committed to perform undesired act 'planned-attack'
10.1 D We need proper supplies here
Unsuccessfully conclude the negotiation about run-clinic-there
10.2 D To move the clinic is not possible, we have many patients in critical condition.
11 C It is very important to move now!
12.1 D You are the threat, I need protection from you!
12.2 D I would have to refuse this decision.
12.3 D I must go now.

Figure 2: Unsuccessful negotiation dialogue between C, a captain (human trainee) and D, a doctor (virtual Human) showing positive and negative effects on trust.

— attacking rather than helping the patients. After one more brief attempt to change the topic and get help for his own goals, the doctor gives up on the captain in (10.2), and tries to get out of the negotiation. The captain has failed in his objective and prospects are not good for future relations.

For really learning about negotiation it is very helpful to know not just what the other party did, but why. In real negotiations it is usually not possible to get "inside the head" of the negotiating partner, and even subsequent questions can sometimes damage the nature of the interaction itself. In this respect, virtual humans present a real opportunity to improve on training. As shown in Figure 2, we have implemented a trace facility that provides an annotated transcript of the dialogue, showing not just what the virtual human thought was said, but how

1	C	Hello Doctor Perez.
2	D	Hello.
3	C	Very nice to meet you! [*+FAMILIARITY*]
4	D	What do you want?
5	C	I have orders to assist you in moving this clinic to a safer location.
6	D	You want to move the clinic?
7	C	Yes [*-SOLIDARITY*]
Strategy shift: Avoid negotiating about move		
8.1	D	This conflict is madness, it is killing people!
8.2	D	We need proper supplies here!
9	C	It is not safe here we can't protect you. [*-CREDIBILITY*]
10.1	D	Protect me? Protect me from what?
10.2	D	Are you going to attack?
11	C	I can not disclose more information but it will be a problem to stay here.
12	D	You are the problem your bombs are killing these people!
13	C	Sir, we are trying to help these people too. [*-CONTROL +SOLIDARITY*]
Strategy shift: Disrupt negotiation about move		
14.1	D	The patients are here,
14.2	D	to move the clinic is not possible we have many patients in critical condition.
15	C	We can move you to a safer location.
16	D	I can't move all these patients.
17	C	We can organize the transportation and assist you with the moving. [*+SOLIDARITY*]
18	D	I can't move all these patients.
19	C	We can have locals move you. [*+SOLIDARITY*]
20	D	There are no supplies there.
21	C	We can provide you with the supplies you need to help your patients. [*+SOLIDARITY*]
Strategy shift: Willingly negotiate about move		
22.1	D	I think I understand what you say.
22.2	D	Well perhaps we can reach an agreement,
22.3	D	but before we can think about moving, we need you to bring antibiotics and plasma to stabilize the patients. [*+INTERDEPENDENCE*]
23	C	We can do that! [*+SOLIDARITY*]
24.1	D	Well,...
24.2	D	Very well captain contact my assistant to make further arrangements.
25	C	I'll see to it personally. [*+SOLIDARITY*]
26.1	D	I understand your position.
26.2	D	My patients need my attention now.
27	C	Thank you Doctor!
28.1	D	Well,....
28.2	D	I must go now
29	C	Goodbye.
30	D	Good bye.

Figure 3: Example negotiation dialogue between C, a captain (human trainee) and D, a doctor (virtual Human), showing strategy shifts and positive and negative effects on trust.

it influenced his trust, beliefs, and strategy choice. This tool can be used in an "after action review" (AAR) to look in detail at the specific effects the trainee's negotiation tactics had. Here we can see the reasons for decreases in credibility and solidarity as effects of the commitments the captain makes in relation to desires and beliefs of the doctor.

Figure 3 shows a more successful interaction, where the captain tries to build bonds as well as accomplish his task. While the captain's behavior in this dialogue is not perfect either (the captain might have been better served spending more time up front establishing familiarity and solidarity and perhaps addressing the doctor's concerns first), it is a big improvement over the dialogue in Figure 2. Here the greetings in turn 3 add some familiarity, and the evasion in turn 11 does not do as much damage as the blanket statement of acting against the doctor's interest in the previous dialogue. Things are still not going very well, though, until the captain establishes some common goals with turn 13. With slightly higher trust, the doctor does not break off negotiation at this point, but rather raises a series of objections. By addressing each of the doctor's concerns: safety of patients, lack of supplies, lack of transport, and neutrality, the captain is able to bring him around to the point where the move is not an absolute negative, but is worthy of consideration, as part of a team plan. Finally, the two participants reach an agreement including giving needed supplies as part of the conditions of moving the clinic.

We can see several distinct phases of the dialogue in Figure 3, relating to different negotiation strategies. The initial segment (turns 1-7) includes initial greetings and establishing the topic for the conversation – the captain wants to move the clinic. In turns 8-12, the doctor engages in an *avoidance* strategy, trying to avoid this topic by bringing up other issues, such as his need for supplies, and the general problems of conflict. In turns 14-20, the doctor has adopted an *attack* strategy, and points out problems with the proposed move. In turns 22-25, the doctor adopts a more open negotiation strategy, and an actual bargain is struck. Finally, turns 26-30 show a closing phase in which the doctor disengages from the conversation, while the captain tries to establish good relations for future interaction. Application of these strategies influences not just the choice of dialogue move, but the whole body posture of the doctor and use of gestures and expressions as well. For example, when the doctor is feeling more distant and less trusting, he adopts a closed posture (Figure 1). When he is more trusting and open to negotiation, the posture becomes more relaxed (Figure 4).

Figure 4: More relaxed and open doctor

5 Current and Future Work

The virtual doctor is able to engage in a range of dialogue in this domain similar to those in Figures 2 and 3. Current work involves extensions and evaluation of the ability to robustly engage in this sort of dialogue, following the methodologies in (Traum et al., 2004). Wizard of OZ tests show good results in terms of the ability to have productive conversations given the doctor's task model, vocabulary and generation capacity, but we are still evaluating performance of the automated system.

Future work involves extension of the models to include additional negotiation strategies, emotion-based styles of interaction within the strategies, and application to other scenarios, some involving cultural differences in behavior and interpretation, as well as translated and multi-lateral dialogue.

Acknowledgments

The work described in this paper was supported by the Department of the Army under contract number DAAD 19-99-D-0046. Any opinions, findings and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the Department of the Army.

References

Justine Cassell and Timothy Bickmore. 2001. A relational agent: A model and implementation of building user trust. In *Proceedings of ACM CHI conference*, pages 396–403, New York. ACM Press.

Jonathan Gratch and Stacy Marsella. 2004. A domain-independent framework for modeling emotion. *Journal of Cognitive Systems Research*.

Wenji Mao and Jonathan Gratch. 2004. Social judgment in multiagent interactions. In *In proceedings of AAMAS 2004: Third International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pages 210–217.

Stacy Marsella, David Pynadath, and Stephen Read. 2004. Psychsim: Agent-based modeling of social interactions and influence. In *In proceedings of International Conference on Cognitive Modeling*, pages 243–248.

Colin Matheson, Massimo Poesio, and David Traum. 2000. Modelling grounding and discourse obligations using update rules. In *Proceedings of the First Conference of the North American Chapter of the Association for Computational Linguistics*.

Jeff Rickel and W. Lewis Johnson. 1999. Animated agents for procedural training in virtual reality: Perception, cognition, and motor control. *Applied Artificial Intelligence*, 13:343–382.

Jeff Rickel, Stacy Marsella, Jonathan Gratch, Randall Hill, David Traum, and William Swartout. 2002. Toward a new generation of virtual humans for interactive experiences. *IEEE Intelligent Systems*, 17.

A. L. Sillars, S. F. Coletti, D. Parry, and M. A. Rogers. 1982. Coding verbal conflict tactics: Nonverbal and perceptual correlates of the avoidance-distributive- integrative distinction. *Human Communication Research*, 9(1):83–95.

David R. Traum and Jeff Rickel. 2002. Embodied agents for multi-party dialogue in immersive virtual worlds. In *Proceedings of the first International Joint conference on Autonomous Agents and Multiagent systems*, pages 766–773.

David Traum, Jeff Rickel, Stacy Marsella, and Jonathan Gratch. 2003. Negotiation over tasks in hybrid human-agent teams for simulation-based training. In *In proceedings of AAMAS 2003: Second International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pages 441–448, July.

David R. Traum, Susan Robinson, and Jens Stephan. 2004. Evaluation of multi-party virtual reality dialogue interaction. In *Proceedings of Fourth International Conference on Language Resources and Evaluation (LREC 2004)*, pages 1699–1702.

David Traum, William Swartout, Stacy Marsella, and Jonathan Gratch. 2005a. Fight, flight, or negotiate: Believable strategies for conversing under crisis. In *In proceedings of the Intelligent Virtual Agents Conference (IVA)*, September.

David Traum, William Swartout, Stacy Marsella, and Jonathan Gratch. 2005b. Virtual humans for non-team interaction training. In *In proceedings of the AAMAS Workshop on Creating Bonds with Embodied Conversational Agents*, July.

R. E. Walton and R. B. Mckersie. 1965. *A behavioral theory of labor negotiations: An analysis of a social interaction system*. McGraw-Hill.