

# Toward Social-Emotional Virtual Humans

## (Doctoral Consortium)

Johnathan Mell  
Supervisor: Jonathan Gratch  
USC Institute for Creative Technologies  
12015 Waterfront Drive  
Los Angeles, CA, USA  
mell@ict.usc.edu

### ABSTRACT

Emotive virtual agents are seen to be a valuable tool in various research domains, including human studies, training, entertainment, and medicine. However, systems that primarily focus on social-emotional agents are largely domain-focused, or require ample customization to make them usable. The importance of agents that display social emotions and behaviors is underscored by the use of social strategies such as favors and ledgers in negotiation contexts, as well as in other domains. A set of core design principles for the next generation of virtual agents is proposed that will result in more robust and varied virtual agents. This early-stage extended abstract serves to illustrate the core principles and continuing research efforts regarding social virtual agents.

### Categories and Subject Descriptors

I.2.11 [Artificial Intelligence]: Human-agent Interaction – *virtual humans*

**Keywords:** Human-agent interaction; Competitions among agents and humans

### 1. Why emotional virtual agents?

Virtual agents are widely used in many tasks, from mediated negotiation tasks [7] to elaborate training systems. Designing virtual humans is a multidisciplinary task, involving not only classical problems of classification and machine learning, but also myriad topics in graphics, animation, natural language processing, and cognitive science. However, at the core of many applications of virtual humans is a necessity to have a clear behavioral model that can be informed by the various input channels and produce believable output behaviors on the virtual character. While there is no strict requirement that the behavior that virtual humans demonstrate be completely identical to that of real humans, systems that are informed by data collected from behavioral studies on humans have shown to be both believable and effective [6]. However, there are many situations, even outside the realm of entertainment, where virtual agents may actually be superior to humans in the same context, such as in the elicitation of patient information [10].

One avenue of research that seems particularly promising is the analysis of virtual humans as training agents that provide feedback on various tasks, including negotiation and conflict resolution. While there is ample evidence that the framing of partners as virtual agent or human has a strong effect in vari-

**Appears in:** *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015),* Bordini, Elkind, Weiss, Yolum (eds.), May, 4–8, 2015, Istanbul, Turkey. Copyright © 2015, International Foundation for Autonomous Agents and Multiagent Systems ([www.ifaamas.org](http://www.ifaamas.org)). All rights reserved.

ous economic games and negotiations [4], there is relatively little work regarding their effectiveness as virtual coaches. Some work has been done on providing simulations of crowds as a dynamic feedback mechanism during public speaking [1], and we wish to extend this work to analyze virtual human coaches in a one-on-one context. Virtual humans are often shown to have less affect than a similar virtual avatar that is controlled by a human [11], but this seeming limitation may actually be beneficial, as it allows them to provide feedback in a more direct and clear manner without risking potential social consequences. As an example that builds on the politeness theory work of Brown and Levinson [2], one future aim is to build virtual humans that can provide feedback that follows various politeness strategies, with the expectations that virtual humans will be able to provide certain types of feedback in a superior way to human tutors.

### 2. Why social emotions?

Though there is profound debate regarding the number and classification of emotions, emotions are accepted as having a profound influence upon human behavior. In addition to allowing predictions to be made concerning the intentions of intelligent actors, understanding and appropriately modeling emotions allows more realistic computational agents to be built that more closely mimic actual human behavior. Furthermore, appraisal theory allows inferences to be made regarding general mood, as well as specific emotional state.

Current computational models, while powerful in analyzing intrapersonal factors, largely do not account for emotional states that rely on *interpersonal* factors. The ‘social emotions’ such as embarrassment, shame, and guilt (and to a lesser extent, mirth and pity) rely on an analysis of the interactions between multiple actors. It is well known that interpersonal factors such as smiles do indeed affect emotional expression, as was demonstrated experimentally [9].

Indeed, not all agents in a given social situation are treated similarly. It has been shown that amusement is more often displayed in social situations when paired with a partner who is a friend rather than a stranger [13]. The important role of social context has likewise been examined; the status of relationships between parties is, perhaps unsurprisingly, important in determining the strength and nature of the parties’ exhibited emotions [14].

To lend further credence to the importance of social emotions, it is important to analyze their purposes as well as their causes. Parkinson argues that one function may be the communication of core relational themes: in the case of guilt, one agent is trying to communicate the idea of self-blame to the other party [15]. If this is interpreted as evidence that the

action preceding this display is less likely to be repeated, then the display of guilt has obvious game theoretic ramifications. Furthermore, such signals have been shown to affect relationship factors such as trust, as well as physiological responses [5].

### 3. Systems that use social emotional agents

The immediate motivation behind this work is to inform the design of virtual humans for teaching negotiation skills, by validating the effectiveness of human negotiation tactics in a virtual context. Virtual humans have shown promise for teaching a variety of interpersonal skills [8]. An important aspect of such teaching is deliberate practice, usually with human role-players. Virtual humans can augment this training by serving as role-playing partners that are infinitely patient, always consistent with pedagogical principles, and able to explain their behavior in terms of course lessons. Some research has already explored the potential of virtual humans for negotiation training [1][3], and work has already been completed that builds upon this body of research and extends its applicability to situations where negotiations repeat over time. These negotiating agents incorporate more complex forms of signaling, such as natural-language dialog, [12] and sometimes involve sophisticated virtual embodiment [3]. Although some research has sought to provide a foundation for using these richer communication channels (e.g., [4] provides a framework for emotional signaling), most of this research has focused on short-term interactions and single negotiations. For a full treatment of the subject of Pareto-efficient over time negotiations, see the author's full paper in this year's AAMAS proceedings.

Due to the importance of displaying robust agents with social emotions as well as the necessity of testing them with real humans, future systems should make use of the following design goals:

- Human-informed – while virtual humans need not act just as humans do, their behavior should be informed by actual behavioral studies using humans.
- Generalizable to a domain – making generic agents within a domain (negotiating agents, conversational agents, question/answer agents) is a key goal.
- Easily customizable – being able to quickly adjust dialogue, goals, and even personality should be a priority.
- Relationship-focused – social-emotional agents will benefit from remembering the results of past encounters.

### 4. Future work

While in traditional thought the feature recognition problem is often seen as entirely separate from the feature generation problem, in virtual agent design this line appears to blur. Designing an agent that recognizes favor-exchange in negotiation leads logically to agents that can exchange favors themselves. Likewise, the need for agents that can give feedback using different politeness face-saving techniques informs the design of natural language understanding systems that can recognize politeness. When informed by actual human-backed data, the feature selection process can become more efficient. The construction of the proposed social-emotional agents is thus anticipated to follow a three-step process.

1. Human studies involving a human participant and a virtual character controlled through a "Wizard of Oz" sys-

tem by a researcher will be used to experimentally validate features of interest, such as favor exchange or politeness.

2. Virtual agents will be designed to generate a wide range of these social-emotional behaviors.
3. The virtual agents will be validated against further human studies that verify that these behaviors are relevant to the context-specific goals.

## REFERENCES

- [1] Batrinca, L., Stratou, G., Shapiro, A., Morency, L.-P., and Scherer, S. (2013). Cicero—towards a multimodal virtual audience platform for public speaking training. In Intelligent Virtual Agents. Springer, 116-128.
- [2] Brown, P. (1987). Politeness: Some universals in language usage (Vol. 4). Cambridge University Press.
- [3] DeVault, D., Mell, J., and Gratch, J. (2015). Toward Natural Turn-Taking in a Virtual Human Negotiation Agent. In AAAI Spring Symposium on Turn-taking and Coordination in Human-Machine Interaction. AAAI Press, Stanford, CA.
- [4] de Melo, C. M., Carnevale, P. J., Read, S. J., & Gratch, J. (2014). Reading people's minds from emotion expressions in interdependent decision making. Journal of personality and social psychology, 106(1), 73.
- [5] Fulmer, A., & Gelfand, M. J. (2010, June). Dynamic trust processes: Trust dissolution and restoration. In IACM 23rd Annual Conference Paper.
- [6] Gratch, J., Ricket, J., André, E., Cassell, J., Petajan, E., & Badler, N. (2002). Creating interactive virtual humans: Some assembly required. University of Southern California, Institute for Creative Technologies, Marina del Rey, CA.
- [7] Guttman, R. H., & Maes, P. (1999). Agent-mediated integrative negotiation for retail electronic commerce. In Agent Mediated Electronic Commerce (pp. 70-90). Springer Berlin Heidelberg.
- [8] Johnsen, K., Raji, A., Stevens, A., Lind, D. S., & Lok, B. (2007, April). The validity of a virtual human experience for interpersonal skills education. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 1049-1058). ACM.
- [9] Kraut, R. E., & Johnston, R. E. (1979). Social and emotional messages of smiling: An ethological approach. Journal of personality and social psychology, 37(9), 1539.
- [10] Lucas, G. M., Gratch, J., King, A., & Morency, L. P. (2014). It's only a computer: virtual humans increase willingness to disclose. Computers in Human Behavior, 37, 94-100.
- [11] Lim, S., & Reeves, B. (2010). Computer agents versus avatars: Responses to interactive game characters controlled by a computer or other player. International Journal of Human-Computer Studies, 68(1), 57-68.
- [12] Rosenfeld, A., Zuckerman, I., Segal-Halevi, E., Drein, O., & Kraus, S. (2014, May). NegoChat: a chat-based negotiation agent. In Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems (pp. 525-532). International Foundation for Autonomous Agents and Multiagent Systems.
- [13] Chapman, A. J. (1983). Humor and laughter in social interaction and some implications for humor research. In *Handbook of humor research* (pp. 135-157). Springer New York.
- [14] Lazarus, R. S. (1991). Progress on a cognitive-motivational-relational theory of emotion. *American psychologist*, 46(8), 819.
- [15] Parkinson, B. (1996). Emotions are social. *British journal of psychology*, 87(4), 663-683.