Niki and Julie: A Robot and Virtual Human for Studying Multimodal Social Interaction

Ron Artstein, David Traum, Jill Boberg,
Alesia Gainer, Jonathan Gratch,
Emmanuel Johnson, Anton Leuski
USC Institute for Creative Technologies
12015 Waterfront Dr., Playa Vista CA 90094
USA
{lastname|ejohnson}@ict.usc.edu

Mikio Nakano
Honda Research Institute Japan
8-1 Honcho, Wako, Saitama 351-0188 Japan
nakano@jp.honda-ri.com

ABSTRACT

We demonstrate two agents, a robot and a virtual human, which can be used for studying factors that impact social influence. The agents engage in dialogue scenarios that build familiarity, share information, and attempt to influence a human participant. The scenarios are variants of the classical "survival task," where members of a team rank the importance of a number of items (e.g., items that might help one survive a crash in the desert). These are ranked individually and then re-ranked following a team discussion, and the difference in ranking provides an objective measure of social influence. Survival tasks have been used in psychology, virtual human research, and human-robot interaction. Our agents are operated in a "Wizard-of-Oz" fashion, where a hidden human operator chooses the agents' dialogue actions while interacting with an experiment participant.

CCS Concepts

•Human-centered computing → Human computer interaction (HCI); Collaborative and social computing; •Computing methodologies → Intelligent agents; •Computer systems organization → Robotics;

Keywords

Human-robot/agent multimodal interaction, Multimodal interactive applications, Affective Computing and interaction

1. INTRODUCTION

One factor that impacts social influence in general, and the survival task in particular, is embodiment. Psychological and communication studies suggest that embodiment increases social influence. For example, participants are less persuaded in a lunar survival task when communicating via teleconference compared with face-to-face interaction [7]. Findings in human-machine interaction are mixed, showing increases in subjective engagement but failing to demonstrate objective persuasion [1]. This is the rationale for using two agents with radically different embodiments.

Another factor that shapes social influence in general, and survival tasks in particular, is the familiarity team members have with each other. In general, people are less influenced by strangers or people they feel more distant from [3]. Telling jokes with machine teammates has been shown to increase persuasion in a lunar survival task [6], and various rapport-building techniques enhance human-machine teamwork [4]. For this reason, the agents also engage in an icebreaker to establish familiarity with the participant.

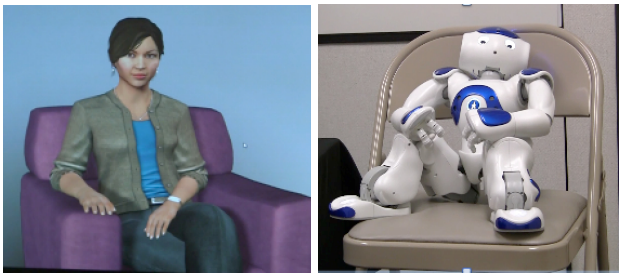

Figure 1: Artificial Agents Julie and Niki

2. SYSTEM DETAILS

We will demonstrate two embodied agents (Figure 1): a NAO robot, named Niki, and a virtual human, named Julie, whose animated body can appear on a monitor. Julie is presented in two modes, either multimodally with voice and virtual embodiment, or voice only accompanied by a static image, as if through a teleconference.

Both agents are controlled by a "Wizard of Oz" system, with a human wizard using a push-button GUI (Figure 2). The interface runs in a web browser and sends messages using the VHMsg messaging protocol to trigger agent behaviors.[1] The system architecture is shown in Figure 3.

Behaviors of Julie, the virtual human, were created using the Virtual Human Toolkit [5], and her voice was synthesized using a voice from NEOspeech's text to speech engine. Niki is a NAO Robot, a humanoid robot commonly used in human robot interaction studies [2]. Niki's non-verbal behaviors were authored using Choregraphe, a multi platform

[1]https://sourceforge.net/projects/vhmsg/

ICMI'16, November 12–16, 2016, Tokyo, Japan
ACM. 978-1-4503-4556-9/16/11...$15.00
http://dx.doi.org/10.1145/2993148.2998532

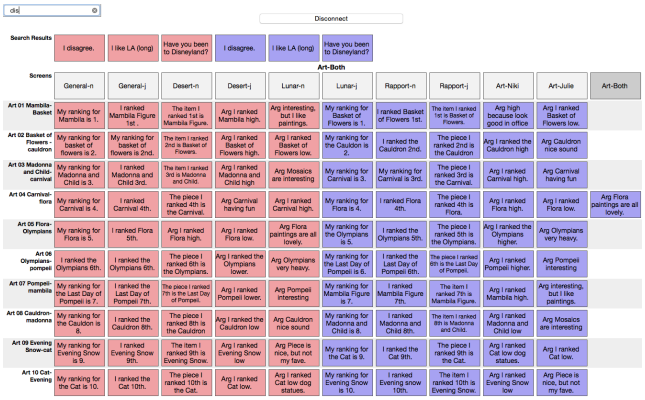

Figure 2: Wizard control interface. The buttons in the second row are used to change screens; the remaining buttons play individual utterances. The top row of buttons is the output of the text search in the top left corner. Red buttons control Niki, Blue control Julie. The label on each button is a mnemonic; the full utterance text appears as a tooltip.

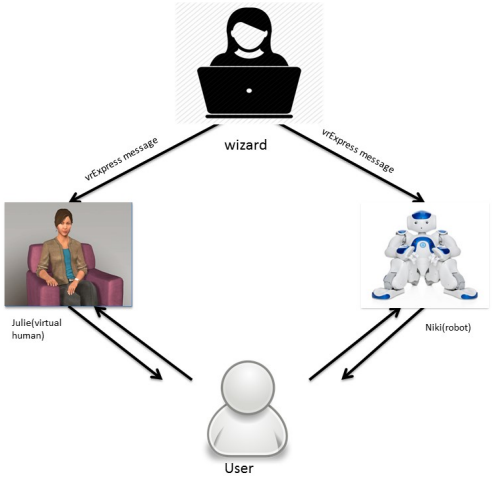

Figure 3: System Architecture

application which allows users to create complex behavior for the NAO robot, and his speech is synthesized using the NAO robot's standard on-board text to speech engine.

3. DEMONSTRATION OUTLINE

Participants will engage in dialogue with one or both of the agents (Figure 4). The demonstration will focus on an original "Save the Art" task, a three-party ranking task between a participant and both agents. Additional dialogues are possible with each agent separately: an ice-breaker designed to create familiarity with each agent, and two classical ranking tasks designed to measure social influence (desert survival and lunar survival).

The ranking tasks ask participants to rank specific items according to their importance, either in terms of their usefulness for survival after landing on the moon or in the desert, or according to the order in which pieces of art should be saved from a fire in an art museum ("Save the Art"). After

Figure 4: Person interacting with the agents

initially ranking the items, the participant and the agents discuss their rankings, and then the participant re-ranks the items; for Save the Art the discussion is between all three parties, with the agents also talking and discussing with each other in the presence of the participant.

The ice-breaker dialogue involves a semi-structured conversation between the agent and the participant. After exchanging greetings, the agent asks the participant a series of seven open-ended questions, such as, "Where are you from?", "What is your favorite kind of music or favorite music artist?", and "Have you traveled?". The agent comments on the participant's answers, and reveals short anecdotes about themselves on the same subjects. This is the only dialogue in which the virtual human is animated; in the ranking tasks, the virtual human is displayed as a static image on a the screen, and only her voice is heard.

4. REFERENCES

[1] S. O. Adalgeirsson and C. Breazeal. Mebot: A robotic platform for socially embodied presence. In *Proc. HRI*, pages 15–22. IEEE Press, 2010.

[2] P. F. Dominey, T. van der Zant, S. Lallee, A.-L. Jouen, X. Hinaut, A. Weitzenfeld, H. van Hoof, and J. Dávila-Chacón. Cooperative human robot interaction with the Nao humanoid: Technical description paper for the "radical dudes". In *RoboCup@Home Technical Description Papers: Germany - Singapore*, 2010.

[3] A. Edmondson. Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44(2):350–383, 1999.

[4] J. Gratch, N. Wang, J. Gerten, E. Fast, and R. Duffy. Creating rapport with virtual agents. In *Proc. IVA 2007*, pages 125–138. Springer, September 2007.

[5] A. Hartholt, D. Traum, S. C. Marsella, A. Shapiro, G. Stratou, A. Leuski, L.-P. Morency, and J. Gratch. All together now: Introducing the Virtual Human Toolkit. In *Proc. IVA 2013*, pages 368–381. Springer, August 2013.

[6] P. Khooshabeh, C. McCall, S. Gandhe, J. Gratch, and J. Blascovich. Does it matter if a computer jokes? In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, pages 77–86. ACM, 2011.

[7] P. J. Werkhoven, J. M. Schraagen, and P. A. J. Punte. Seeing is believing: communication performance under isotropic teleconferencing conditions. *Displays*, 22(4):137–149, 2001.