**SIGdial 2018: Special Session on Physically Situated Dialogue**

**Topic Summary**
Recent technologies have brought conversational robots out of the lab and into the homes and workplaces of real users. Dialogue is now actively taking place between people and agents (e.g., robots, devices) to operate, navigate, and manipulate physical space. *Physically situated dialogue* distinguishes itself from other forms of dialogue in that it (1) takes place in a physical space, (2) refers to the shared surroundings of interlocutors, and (3) involves a physical agent (robot or device). While dialogue offers many potential benefits, such as ease of use and hands-free operation, understanding the nature of effective physically situated dialogue—encompassing turn-taking, grounding, and operating with limited amounts of data—will be essential. We believe the timing is right for a SIGdial special session highlighting this topic.

At the same time, there is ongoing, exciting growth with interdisciplinary efforts to bring together natural language processing and dialogue with the fields of robotics, human-robot interaction, and computer vision. This is evidenced by recent workshops such as the 2017 ACL Workshop on Language Grounding for Robotics (RoboNLP), the 2017 AAAI Fall Symposium on Natural Communication for Human-Robot Collaboration, and the 2017 NIPS Workshop on Visually-Grounded Interaction and Language. Meetings such as the special session we propose enable the communication necessary to share progress across disciplines. Moreover, there is a growing need for bi-directional dialogue work to be showcased that supports the body of work on language grounding and vision and language, as well as dialogue that will allow physically situated agents, like robots, to ask for clarification and provide updates on internal state (e.g., common ground).

**Topic Objectives**
Given the success of 2017 workshops focused on language, robots, and vision, we propose this special session as a venue to continue the shared conversation of the highly interdisciplinary field of physically situated dialogue. Proposal organizers attended one or more of the above workshops and believe that a SIGdial special session would be a good venue to continue to bring the communities of robotics, computer vision, natural language processing, and dialogue together. A SIGdial special session would provide a welcoming way for some of the previous participants of the workshops to better understand the crucial role dialogue must play in language grounding, and more generally, human-robot communication.

**Topic Areas including, but not limited to:**
- Interaction studies with smart-home devices
- Learning from demonstration through natural language dialogue
- Explainable AI in physical spaces
- Representations of physical surroundings / world modeling to support grounded communication
- Embodied visual question answering and/or generation
- Empirical studies of human-robot dialogue (Wizard-of-Oz based, simulated, semi-autonomous)
- Computational models of dialogue management and/or turn-taking with physical agents
- Methods of building common ground with physical agents in real-world or simulated environments
- Corpora of physically situated dialogue (Wizard-of-Oz based or otherwise)
- Multimodal information processing to support dialogue (including speech, gaze, gesture)
- Physical embodiment, voice, or personification of robots and its effect on human-robot dialogue
- Communicating feedback from robots using affordances in addition to speech
- Spoken language generation for physically situated dialogue

**List of Organizers**
- Sean Andrist, researcher at Microsoft Research in the Perception and Interaction group. His research interests involve designing, building, and evaluating socially interactive technologies that are physically situated in the open world. He received his PhD from the Department of Computer Sciences at the University of Wisconsin–Madison.
- Stephanie Lukin, Computer Scientist at the Army Research Lab. Her research interests include natural language generation, customization, and personalization of human interactions with virtual agents or robots. She received her Ph.D. and M.S. degrees from the Natural Language and Dialogue Systems Lab at the University of California, Santa Cruz.
- Matthew Marge, Research Scientist at the Army Research Lab. **(Primary contact)** matthew.r.marge.civ@mail.mil. His current interests lie at the intersection of computational linguistics and human-robot interaction, specializing in designing and building dialogue systems for robots. He received the Ph.D. and M.S. degrees in Language and Information Technologies from the School of Computer Science at Carnegie Mellon University.
- Jesse Thomason, Graduate student completing a PhD in 2018, University of Texas at Austin. His research interests are in semantic understanding and language grounding in human-robot conversations. He focuses on expanding embodied robot language understanding through dialog with humans.
- Zhou Yu, Assistant Professor at Computer Science Department at the University of California, Davis. She received her PhD in 2017 from Carnegie Mellon University. Dr. Yu's research focuses on designing algorithms for real-time multimodal intelligent interactive systems that coordinate with user actions. She optimizes human-machine communication via studies of multimodal sensing and analysis, speech and natural language processing, machine learning and human-computer interaction.

**List of Potential Participants**
- Robotics Experts: Sonia Chernova (GA Tech), Thomas Howard (Rochester), Ross Knepper (Cornell), Jean Oh (CMU), Nick Roy (MIT), Jivko Sinapov (Tufts), Siddhartha Srinivasa (UW), Stefanie Tellex (Brown), Manuela Veloso (CMU), Matthew Walter (TTIC)
- Vision Experts: Mohit Bansal (UNC Chapel Hill), Dhruv Batra (GA Tech/FAIR), Abhishek Das (GA Tech), Devi Parikh (GA Tech/FAIR)
- HRI Experts: Daniel Barber (UCF), Maya Cakmak (UW), Laura Hiatt (NRL), Susan G. Hill (ARL), Malte Jung (Cornell), Jill Lehman (Disney Research), Iolanda Leite (KTH), Bilge Mutlu (UW Madison), Matthias Schuetz (Tufts), Tom Williams (Colorado School of Mines)
- NLP Experts: Yoav Artzi (Cornell), Yonaton Bisk (UW), Dan Bohus (MSR), Yejin Choi (UW)
- Dialogue Experts: Joyce Chai (MSU), Mary Ellen Foster (Glasgow University), Ryuichiro Higashinaka (NTT Media Lab), Oliver Lemon (Herriott-Watt), Mikio Nakano (HRI), Alexander Rudnicky (CMU), Gabriel Skantze (KTH), David Traum (USC), Marilyn Walker (UCSC)
- Additionally, RoboNLP participants (https://robonlp2017.github.io) and AAAI FSS NCHRC participants (http://www.ttic.edu/nchrc)

**Requested Format:**
A 2.5 hour session showcasing one **oral paper session** (full SIGdial papers), a **poster session** (or short talks) with position papers (2-4 page submissions) highlighting work in progress that shows the integrative nature of the topic, and a **panel discussion** featuring a mix of experts in robotics, dialogue, and computer vision, continuing the conversation from previous panels at ACL RoboNLP and the AAAI Fall Symposium.