DEFINING NEXT-GENERATION MULTI-MODAL COMMUNICATION IN HUMAN ROBOT INTERACTION

Stephanie Lackey, Daniel Barber, Lauren Reinerman
University of Central Florida

Norman I. Badler
University of Pennsylvania

Irwin Hudson
Army Research Laboratory

With teleoperation being the contemporary standard for Human Robot Interaction (HRI), research into multi-modal communication (MMC) has focused on development of advanced Operator Control Units (OCU) supporting control of one or more robots. However, with advances being made to improve the perception, intelligence, and mobility of robots, a need exists to revolutionize the ways in which Soldiers interact with robotic team members. Within this future vision, mixed-initiative Soldier-Robot (SR) teams will work collaboratively sharing information back-and-forth in a fluid natural manner using combinations of communication methods. Therefore, new definitions are required to focus research efforts to support next-generation MMC. After a thorough survey of the literature and a scientific workshop on the topic, this paper aims to operationally define MMC, Explicit Communication, and Implicit Communication to encompass the shifting paradigm of HRI from a controller/controlled relationship to a cooperative team mate relationship. This paper presents the results from a survey of the literature and a scientific workshop that inform proposed definitions for multi-modal, explicit, and implicit communication. An illustrative scenario vignette provides context and specific examples of each communication type. Finally, future research efforts are summarized.

Background

Typical military operations that involve robot team mates focus on search and rescue, ordinance disposal, mine clearing, and remote targeting missions. Existing interfaces emphasize teleoperation because robots function as tools for their human operators. This practice illustrates an evolutionary step on the path to autonomous machine intention (Redden and Elliot, 2010; Hancock, 1996), but lacks the level of interaction required for human-robot team collaboration. Achieving Soldier-Robot (SR) team collaboration requires integration of concepts and products from diverse technical areas including:
perception, intelligence, human-robot interaction (HRI), dexterous manipulation, and unique mobility. An interdisciplinary approach sets the stage for advancing the science of human-robot teaming by delivering highly capable, autonomous robotic systems. Concretely, this means a paradigm shift from controlling robots to collaborating with them, operating robots as tools to viewing them as team members, and evolving from display-centered interfaces to natural interaction.

Thus, a need exists to revolutionize the ways in which Soldiers interact with robot team members. Three critical terms require revision in order to advance Soldier-Robot Interaction (SRI): Multi-Modal Communication (MMC), Explicit Communication, and Implicit Communication.

Operational Definitions

HRI research investigating MMC presents an opportunity to address the SRI paradigm shift described above. Recognizing a need to evolve the definition of key terms due to the complex, noisy, and chaotic nature of the dismounted infantry operational environment, this research effort surveyed the literature in order to identify theoretical foundations to support basic and applied research under the Army Research Laboratory's (ARL) Robotics Collaborative Technology Alliance (RCTA). A scientific workshop focused on HRI communication issues followed in December 2010. The results of the workshop influenced development of operational definitions for MMC, explicit communication , and implicit communication . Each term is defined below and illustrative examples describing the incorporation of explicit and implicit communication follow.

Multi modal communication. A blend of auditory, visual, and tactile modes is anticipated to be of most benefit to mixed-initiative teams in the operational environment. It is important to consider modality types and delivery methods, and to differentiate between explicit and implicit modes. The items listed in Table 1 represent modalities and methods that meet the criteria of the operational constraints and reflect findings in the literature filtered by the operational constraints of military environment. Specifically, the content of Table 1 summarizes the modalities and delivery methods enabling effective human-human communication.

Table 1. RCTA HRI communication modalities

Modality	Delivery	Explicit	Implicit
Auditory	Speech, Sounds	Language	Tone, Rate, Pitch
Visual	Posture, Facial Expression, Gesture, Gait, Social Distance	Intentional Pointing, Hand Signals	Unintentional Body Language, Intensity, Eye Contact, Talking with Hands, Emotions
Tactile	Belt, Vest	Intentional Touching, Patterns	Pressure, Patterns, Shakiness

Humans naturally communicate with one another using redundant and simultaneous communication modalities that are flexibly switched to meet diverse needs (Oviatt, 2002; Oviatt, Coulston & Lunsford, 2004) and convey meaning and context at multiple levels of complexity (Bischoff & Graefe, 2002). Further, MMC positively impacts effectiveness and efficiency (Parr, 2004; Oviatt, 2000). These features represent critical and significant advancements in SR interaction and team performance. Therefore, the following operational definition is proposed.

Multi-modal communication is the exchange of information through a flexible selection of explicit and implicit modalities that enables interactions and influences behavior, thoughts, and emotions.

Explicit communication. With teleoperation being the contemporary standard for explicit communication from humans to robots, a goal for the December 2010 workshop was to develop a definition of explicit communication for use within the RCTA to support mixed-initiative operational teams. Within this new paradigm, explicit communication must support the purposeful communication of information through auditory, visual, or tactile modalities to influence the behavior, thoughts, and emotions of others. Through an examination of the literature, different methods of communication across modalities were investigated with a focus on bi-directional capabilities. For example, voice commands have shown reduced operation time in discrete robotic tasks, but perform poorly for continuous tasks (Redden, Carstens, Pettitt, 2010). Additionally, auditory cues were shown to increase performance within robotic tasks when used alone and in conjunction with tactile cues (Gunn, Warm, Nelson, Bolia, Schumsky, Corcoran, 2005; Haas, 2007). Arm and hand gestures are a natural and intuitive method of communication between humans (Wexelblat, 1995) and can be quicker and feel more intuitive than manual controls (Guo & Sharlin 2008). Tactile communication from robot to Soldier via skin stimulation is a method under investigation because it is hands-free, stealthy, and shown to increase situational awareness, decrease sensory overload, reduced response time, and even motor task training and reduced response time due to its egocentric directional cueing (White, 2010; Elliott, Coovert, Prewett, Walvord, Saboe, & Johnson, 2009; Hutchins, Cosenzo, McDermott, Feng, Barnes, & Gacy, 2009; Bloomfield & Badler, 2008). Based on this review and operational requirements, the following operational definition for HRI research in explicit communication is proposed.

Explicit Communication is the purposeful conveyance of information through multiple modalities (i.e., audio, visual, tactile) that has a defined meaning.

Implicit Communication. Implicit communication is arguably the most difficult phenomenon to define in the context of HRI. Various definitions exist with differing components of what makes up implicit communication (Bauer, Wollherr, & Buss, 2007; Lin, Le, Becker, & Makedon, 2010; Rani, Sarkar, Smith, & Adams; 2008). The closest synonym for implicit communication is nonverbal communication. Examples of nonverbal communication are body posture, eye contact, touching, and social distance. However, as Richmond, McCroskey, & Hickson (2008) stated, "A difficult issue facing scholars of nonverbal communication has been the drawing of meaningful and clear distinctions between verbal and nonverbal messages." Aside from this challenge, many nonverbal communication taxonomies exist and no two are the same. Not to mention that while nonverbal communication seems to be the closest description of implicit communication, implicit communication sometimes includes verbal communication such as paralinguistics (e.g., grunts, sighs, tone, rate of speech). Mehrabian (1981) addressed some of these issues with his theory of implicit communication. That theory focuses on implicit communication as conveying emotions and attitudes through five major categories: Emblem, Illustrator, Affect Display, Regulator, and Adaptor. Considering the available literature, the following operational definition for HRI research is proposed.

Implicit Communication is the inadvertent conveyance of information about emotional and contextual state that will affect interpretation, thoughts, and behaviors.

Conceptual Illustration: "Follow That Guy"

In order to convey how MMC is used within an operational context and to help differentiate between explicit and implicit communication, a vignette composed of "mission primitives" is used. A vignette is a short impressionistic scene that focuses on one moment or gives an idea about a character, setting, or object. Mission primitives describe specific action/interaction within a vignette and can be combined and reused. Through the identification of mission primitives it is possible to develop MMC solutions to support multiple vignettes and the larger scenarios they support.

In the following example, illustrated by Fig. 1, a Soldier (A) and a robot teammate (B) are performing a reconnaissance and surveillance mission in an urban environment. At some time while traversing the city street, the robot teammate identifies a Point of Interest (POI) (e.g., enemy target) and alerts the Soldier. The Soldier then requests the robot teammate to follow the POI (C), while he/she moves to a Non-Line-of-Site (NLOS) location. As the robot teammate follows the POI, it keeps the Soldier informed of events.

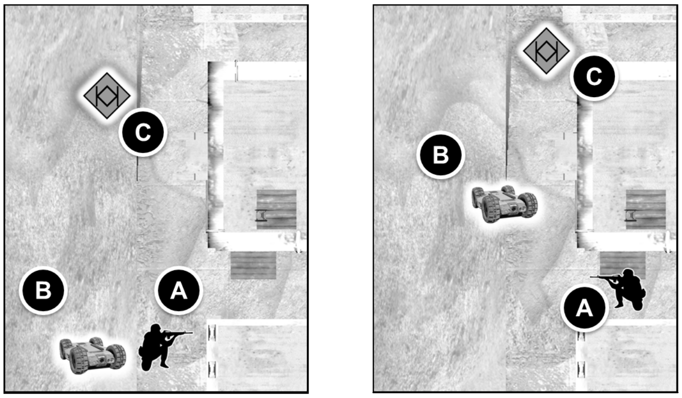

Figure 1 - Point of Interest (POI) vignette

From the start to the end of this vignette, there are several communication events that illustrate mission primitives. In Table 2, each mission primitive is described as a combination of explicit or implicit communication over a selected modality.

Table 2 – POI vignette communication events

Time	Events	Modality/Explicit	Modality/Implicit
T1	Soldier requests robot to follow him/her	Visual: Gesture	Visual: Urgency
T2	Robot acknowledges request	Tactile: Cue	Visual: Urgency
T2B	Robot Follows POI	None	Visual: Social Distance
T3	Robot notifies Soldier of POI	Tactile: Cue & Navigational Aide	Tactile: Context
T4	Soldier request robot to follow POI	Visual: Gesture	Visual: Urgency
T5	Robot acknowledges request	Tactile: Cue	Visual: Urgency
T6	Soldier moves to NLOS location. Robot follows POI.	None	Visual: Social Distance
T7	Robot notifies Soldier of change in direction (e.g. takes a right down street)	Tactile: Cue & Navigational Aide Audio: Cue or Voice	Tactile: Context and Urgency Audio: Context and Urgency

For example, at time T1, the Soldier uses an explicit visual gesture to signal the robot teammate to follow him/her. While performing this gesture the robot teammate looks for additional implicit information in the gesture to identify the urgency of the request. At times T2B and T6, the robot is following the POI, using social distance as a method of implicit communication with the POI. Finally, at time T7, the robot is communicating information to the Soldier via tactile and audio modalities to provide both explicit and implicit information about its change in direction while following the POI. Following mission primitives in this example, it is clearly shown how both modalities and explicit and implicit communications are used to support bi-directional communication within a mixed-initiative team. Moreover, this brief vignette illustrates the importance of concurrent, diverse, and/or switchable communication modes to effectively convey meaning in the operational environment.

Summary

The proliferation of unmanned assets (e.g., robots) within the military services indicates that the next-generation of SR operational protocols will undoubtedly incorporate novel SR team collaboration requirements; thus, necessitating equally innovative SRI communication methods. True integration of robotic assets within SR teams requires the scientific community to rapidly evolve the concept and implementation of human-robot communication for dismounted applications. Future systems will use traditional controllers and other device-based interfaces less, transitioning to natural bi-directional communications.

Given emerging trends in the fields of HRI and MMC, the research community is poised to significantly impact the imminent SRI communications revolution. A natural next step to enabling Soldiers to shift from controlling robots to collaborating with them is to enable Soldiers to communicate with robots in ways that are natural, efficient, and effective, is to reassess the current definitions of MMC, explicit communication, and implicit communication. This paper builds upon existing interdisciplinary research to advance these terms in support of basic and applied research conducted within ARL's RCTA. The definitions proposed for multi-modal, explicit, and implicit communication begin to bridge the gap between the current state-of-the-science and future mixed-initiative team applications. Further research and experimentation is required to test the proposed definitions, facilitate their refinement, and develop new technologies.

RCTA efforts planned for 2011 include basic and applied research in the area of MMC applied to the operational environment. Explicit communication research includes advancing the use of Parameterized Action Representation (PAR) to control simulated robots in an operationally relevant scenario (Pelechano, Allbeck, & Badler, 2008). Tactile encoding schemes for complex messages (e.g., cue and navigational aide) and bidirectional communication will be developed and assessed in live-virtual-constructive environments. Basic research in the area of implicit communication aims to develop foundational methods for detecting, classifying, and demonstrating implicit delivery mechanisms. The initial research conducted in 2011 will serve as a launching point for integration of innovative communication modalities within advanced robotic systems.

Acknowledgement

This work was supported by the US Army Research Laboratory (ARL) (100005MC). 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 ARL or the US Government. The US Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

REFERENCES

Bauer, A., Wollherr, D., & Buss, M. (2007). Human-robot collaboration: A survey. *International Journal of Humanoid Robotics, 12*(8), pp. 1-20.

Bischoff, R. & Graefe, V. (2002). Dependable multimodal communication and interaction with robotic assistants. *Proceedings of the 11th IEEE International Workshop on Robot and Human Interactive Communication*, Berlin, 300-305.

Bloomfield, A & Badler, N. (2008) Virtual training via vibrotactile arrays. *Presence J*., 17(2), 103-120.

Elliott, L. R., Coovert, M. D., Prewett, M., Walvord, A. G., Saboe, K., & Johnson, R. (2009). *A review and meta analysis of vibrotactile and visual information displays* (ARL-TR-4955*)*. Aberdeen Proving Ground, MD: U.S. Army Research Laboratory, 1-36.

Gunn, D. V., Warm, J. S., Nelson, W. T., Bolia, R. S., Schumsky, D. A., & Corcoran, K. J. (2005). Target acquisition with UAVs: Vigilance displays and advanced cueing interfaces. *Human Factors,* 47(3), 488-497.

Guo, C., & Sharlin, E. (2008). Exploring the use of tangible user interfaces for human-robot interaction: A comparative study. *Proceedings of CHI'08, ACM,* 1-10.

Haas, E. C. (2007). Integrating auditory warnings with tactile cues in multimodal displays for challenging environments. *Proceedings of the 13th International Conference on Auditory Displays*, Montreal, Canada, 127-131.

Hancock, P.A. (1996). On convergent technological evolution. *Ergonomics in Design*, 41, 22-9.

Hutchins, S., Cosenzo, K., McDermott, P., Feng, T., Barnes, M., & Gacy, M., (2009). An investigation of the tactile communications channel for robotic control. *Proceedings of Human Factors and Ergonomics Society Annual Meeting,* 53(4), 182-186.

Lin, Y. Y., Le, Z., Becker, E., & Makedon, F. (2010). Acoustical implicit communication in human-robot interaction. *Proceedings of the 3rd International Conference on PErvasive Technologies Related to Assistive Environments (PETRA)*.

Mehrabian, A. (1981). *Silent messages: Implicit communication of emotions and attitudes*. Belmont, CA: Wadsworth Pub. Co.

Nigay, L. & Coutaz, J. (1993). A Design Space for Multimodal Systems: Concurrent Processing and Data Fusion. *In Proceedings of INTERCHI'93*, Amsterdam, Netherlands, 172-178.

Oviatt, S. (2002). Breaking the Robustness Barrier: Recent Progress on the Design of Robust Multimodal Systems. *Advances in Computers*. San Diego: CA. Elsevier Sciences.

Oviatt, S. (2000). Taming Recognition Errors with a Multimodal Interface. *Communications of the ACM*, 43(9), 45-51.

Oviatt, S., Coulston, R., Lunsford, R. (2004). When Do We Interact Multimodally? Cognitive Load and Multimodal Communication Patterns. *Proceedings of the 6th International Conference on Multimodal Interfaces*, New York, NY: ACM Press, 129-136.

Parr, L. A. (2004). Perceptual Biases for Multimodal Cues in Chimpanzee (Pan troglodytes) Affect Recognition. *Animal Cognition*, 7, 171-178.

Pelechano, N., Allbeck, J., & Badler, N. (2008). Virtual Crowds, Methods, Simulation and Control. *Synthesis Lectures on Computer Graphics and Animation*, Morgan & Claypool.

Rani, P., Sarkar, N., Smith, C. A., & Adams, J. A. (2003). Affective communication for implicit human-machine interaction. *IEEE International Conference on Systems, Man and Cybernetics, 2003*(5), 4896-4903.

Redden, E. S., Carstens, C. B., & Pettitt, R. (2010). *Intuitive speech-based robotic control* (ADA519652). Aberdeen Proving Ground, MD: U.S. Army Research Laboratory, 1-50.

Redden, E., & Elliott, L. (2010). Robotic Control Systems for Dismounted Soldiers. In M. J. Barnes & F. G. Jentsch (eds), *Human-Robot Interactions in Future Military Operations* (pp. 335–352). Farnham: Ashgate Publishing.

Richmond, V.P., McCroskey, J.C., & Hickson, M.L. (2008). *Nonverbal Behavior in Interpersonal Relations*. Boston: Allyn & Bacon.

Wexelblat, A. (1995). An approach to natural gesture in virtual environments. *ACM Transactions on Computer-Human Interaction*, *2*(3), 179-200.

White, T. (2010). *Suitable body locations and vibrotactile cueing types for dismounted soldiers*. (ARL-TR-5186*)*. Aberdeen Proving Ground, MD: U.S. Army Research Laboratory, 1-38.