Building Blocks of Social Intelligence:
Enabling Autonomy for Socially Intelligent and Assistive Robots

Ross Mead, Amin Atrash, Edward Kaszubski, Aaron St. Clair,
Jillian Greczek, Caitlyn Clabaugh, Brian Kohan, Maja J Matarić

University of Southern California
{rossmead, atrash, kaszubsk, astclair, greczek, clabaugh, bkohan, mataric}@usc.edu

Introduction

We present an overview of the control, recognition, decision-making, and learning techniques utilized by the Interaction Lab (robotics.usc.edu/interaction) at the University of Southern California (USC) to enable autonomy in sociable and socially assistive robots. These techniques are implemented with two software libraries: 1) the Social Behavior Library (SBL) provides autonomous social behavior controllers; and 2) the Social Interaction Manager (SIM) provides probabilistic models to recognize, reason over, and learn about human behavior. Both libraries are implemented in the Robot Operating System (ROS; www.ros.org) framework, and are made available to the community as open-source software packages in the USC ROS Package Repository (code.google.com/p/usc-interaction-software).

Social Behavior Library (SBL)

The Social Behavior Library (SBL) provides generic computational models of social behavior, which are implemented as autonomous controllers. The following social behavior categories are represented in SBL: proxemics (social spacing), vocalics (nonverbal speech cues), oculesics (eye gaze), kinesics (gesture), deixis (referencing), and chronemics (timing); each is described below.

Proxemics is the study of interagent social spatial behavior. In Mead & Matarić (2014), we model proxemic behavior as an optimization problem that considers the production and perception of social stimuli (e.g., speech and gesture) as a function of position and environmental interference. These factors are represented as a dynamic Bayesian network, over which we sample various positions in the environment to maximize social signal recognition rates for all agents (human and robot) in the interaction.

Vocalics is the study of the nonverbal aspects of speech, such as volume, pitch, and rate. Our contribution is a parametric vocalic behavior controller that autonomously adjusts the robot speaker volume based on models of how a human user will hear speech produced by the robot. These models vary with distance, orientation, and perceived environmental interference (Mead & Matarić 2014). Our future work will investigate adapting the pitch and rate of speech produced by a robot to improve user speech perception.

Oculesics is the study of interagent eye gaze behavior. In SBL, our implementations of oculesic behavior are based on three models: 1) a neurobiological model of the human visual attention system (Itti, Dhavale, & Pighin 2003); 2) a conversational gaze model for multi-party interactions (similar to Mutlu et al. 2012); and 3) a model of robot performance requirements to conduct a functional interaction (based on results from Mead & Matarić 2014). These models are combined in a weighted mixture model to generate autonomous eye gaze behavior that is natural and socially appropriate for the human user, as well as functional for the autonomous sociable robot.

Kinesics is the study of communicative body movement using the hands, arms, torso, head, and face. In Mead et al. (2010), we demonstrated a kinesic behavior controller that selects and executes appropriate arm gestures and head movements to produce in coordination with speech.

Deixis is the study of the social referencing of an entity (often through pointing gestures). Deictic gestures can be used nonverbally, or can coincide with words like "this" and "that", to indicate a referent target in physical space. In St. Clair, Mead, & Matarić (2011), we present three robot deictic behavior control strategies (arm-only, eye/head-only, and both) from which the robot can autonomously select to maximize human perception of a deictic referent.

Chronemics is the study of timing during social interactions. Two fundamental chronemic behaviors—turn-taking and backchanneling (e.g., head nods)—are being integrated into SBL based on the work of Chao & Thomaz (2013) and Morency, de Kok, & Gratch (2009), respectively.

We are in the process of experimentally validating each component of SBL. SBL aims to serve as a tool to robotics researchers, providing fundamental behavior controllers for autonomous sociable robots. We now discuss mechanisms to determine when to use these controllers.

Social Interaction Manager (SIM)

Autonomous decision-making in real-world HRI is challenging due to uncertainty in both sensing and action, incomplete information about the environment, and large state spaces. In the USC Interaction Lab, we are developing the Social Interaction Manager (SIM) to explore approaches to solve a variety of problems faced by socially intelligent and assistive robots in real-world interactions. With SIM, we answer the question of *what* to do, whereas with SBL, we answer the question of *how* to do it.

Recognition. In Fasola & Matarić (2014), we utilize Bayesian methods to reason over spatial cues in spoken dialogue, such as "on", "near", "away from", and "between". We perform statistical grammar analysis over user speech to generate semantic fields relating entities in the environment, mapping language to meaning and action. As an alternative to vocal data, in Mead, Atrash, & Matarić (2013) we utilize hidden Markov models (HMMs) over visual data to predict and recognize transitions into (initiation) and out of (termination) human-robot interactions. We also utilize HMMs in Atrash et al. (2011) to recognize human behavior from physiological data during strenuous activities, which is then used to inform a decision-making mechanism to provide motivation feedback to improve user performance.

Decision-making. In Greczek et al. (2014), we employ Markov Decision Processes (MDPs) to autonomously adjust the level of feedback provided by a robot to a human user engaged in an interaction. We also utilize MDPs in St. Clair & Matarić (2013) to model communication during collaborative human-robot activities; this work focuses on developing effective robot teammates by ensuring that robots account for the actions of other people in their high-level task planning and provide effective feedback in support of people's natural collaboration abilities. These approaches enable a robot to clarify its instructions and intentions using speech and gesture to guide a person through a task, with the goal of improving objective and subjective measures of user performance.

Learning. Decision-making algorithms typically assume an accurate representation of the world; however, this is often an unrealistic assumption. To address this, we are developing a novel Bayesian reinforcement learning algorithm to enable learning of probabilistic graphical model parameters from natural communication with people (St. Clair et al. 2011). The algorithm provides a data-driven approach to replicating a policy provided by a human, taking advantage of error models of recognition systems, and allowing for rapid adaptation when needed. We are also exploring the use of multi-task learning techniques to optimize the parameters for individual users over long-term interactions. Such methods allow us to capitalize on extended individual interactions while still considering commonalities across the population.

Applications

The approaches presented for control in SBL, as well as those for recognition, decision-making, and learning in SIM, are being integrated into our general social interaction software pipeline to provide a suite of autonomous capabilities for socially intelligent and assistive robots.

The various components of SBL and SIM have been highlighted in a variety of real-world domains with both general populations and populations with special needs. A preliminary implementation of SBL was demonstrated in a rehabilitation context with people post-stroke (Mead et al. 2010). MDP-based decision-making was deployed on a NAO robot interacting with children with autism (Grezcek et al. 2014), and collaboration and robot learning have been demonstrated in an augmented environment where a human and robot work together to gather virtual moving objects under constraints (St. Clair & Matarić 2013). Based on these experiences with real-world populations, we believe that the use of these intelligent approaches will help facilitate rich autonomy and the deployment of robots in unstructured and dynamic interactions and environments.

Both SBL and SIM are implemented in the ROS framework, and are made available to the community as open-source software packages in the USC ROS Package Repository (code.google.com/p/usc-interaction-software). Furthermore, these libraries and techniques are being made more accessible through simplified software tools that enable the integration of sociable robotics into K-12 educational curricula to promote studies in science, technology, engineering, and math (Mead, Zeltner, & Matarić, 2013; Lu & Mead, 2012).

Acknowledgments

References

Atrash, A.; Mower, E.; Shams, K.; and Matarić, M. 2011. Recognition of Physiological Data for a Motivational Agent. In *The 2011 AAAI Spring Symposium Series: Computational Physiology*. Stanford, CA.

Chao, C. and Thomaz, A.L. 2013. Controlling Social Dynamics with a Parametrized Model of Floor Regulation. *Journal of Human-Robot Interaction* 2(1):4-29.

Fasola, J. and Matarić, M.J. 2014. Interpreting Instruction Sequences in Spatial Language Discourse with Pragmatics towards Natural Human-Robot Interaction. In *Proceedings of ICRA-2014*, to appear. Hong Kong, China.

Greczek, J.; Kaszubski, E.; Atrash, A.; and Matarić, M.J. 2014. Graded Cueing Feedback in Robot-Mediated Imitation Practice for Children with Autism Spectrum Disorders. In *Proceedings of Ro-Man 2014*, to appear. Edinburgh, Scotland.

Itti, L. and N. Dhavale. 2003. Realistic Avatar Eye and Head Animation Using a Neurobiological Model of Visual Attention. In Proceedings of SPIE 48th Annual International Symposium on Optical Science and Technology, 64-78. Bellingham, WA: SPIE Press.

Lu, D.V. and Mead, R. 2012. Introducing Students Grades 6-12 to Expressive Robotics. In *Proceedings of HRI-2012*, 411-412. Boston, MA.

Mead, R. and Matarić, M.J. 2014. Perceptual Models of Human-Robot Proxemics. In *Proceedings of ISER-2014*, to appear. Marrakech/Essaouira, Morocco.

Mead, R.; Atrash, A.; and Matarić, M. 2013. Automated Proxemic Feature Extraction and Behavior Recognition: Applications in Human-Robot Interaction. *International Journal of Social Robotics* 5(3):367-378.

Mead, R.; Wade, E.R.; Johnson, P.; St. Clair, A.; Chen, S.; and Matarić, M.J. 2010. An Architecture for Rehabilitation Task Practice in Socially Assistive Human-Robot Interaction. In *Proceedings of Ro-Man 2010*, 404-409. Viareggio, Italy.

Mead, R.; Zeltner, S.; and Matarić, M.J. 2013. Integrating ROS into Educational Robotics: Bridging the Gap between Grade School and Grad School. In *Proceedings of GCER-2013*. Norman, OK.

Morency, L.-P.; de Kok, I.; and Gratch, J. 2009. A Probabilistic Multimodal Approach for Predicting Listener Backchannels. *Journal of Autonomous Agents and Multi-Agent Systems* 20(1):70-84.

Mutlu, B.; Kanda, T.; Forlizzi, J.; Hodgins, J.; and Ishiguro, H. 2012. Conversational Gaze Mechanisms for Humanlike Robots. *ACM Transactions on Interactive Intelligent Systems* 1(2):1-33.

St. Clair, A.B. and Matarić, M.J. 2013. Role-Based Coordinating Communication for Effective Human-Robot Task Collaborations. *The 4th International CTS Workshop on Collaborative Robots and Human Robot Interaction*, 241-244. San Diego, CA.

St. Clair, A.; Atrash, A.; Mead, R.; and Matarić, M.J. 2011. Speech, Gesture, and Space: Investigating Explicit and Implicit Communication in Multi-Human Multi-Robot Collaborations. In *The 2011 AAAI Spring Symposium Series: Multi-Robot Systems and Physical Data Structures*. Stanford, CA.

St. Clair, A.B.; Mead, R.; and Matarić, M.J. 2011. Investigating the Effects of Visual Saliency on Deictic Gesture Production by a Humanoid Robot. In *Proceedings of Ro-Man 2011*, 210-216. Atlanta, GA.