POMDP Planning for Socially Situated Tasks in HRI

Frank Broz

The emerging field of study that is human-robot interaction continues one of the main goals of general robotics research: to create robots that can serve as effective helpmates to humans. Building on the work modeling and acting in physical environments that have led to great successes in navigation and other lower level capabilities, HRI takes on an even more difficult challenge: that of modeling and acting with human participants. As a branch of robotics, a focus on tasks that can be empirically evaluated and approaches that are informed by existing robotics and artificial intelligence will assure that HRI produces robots that are functional as well as compelling. However, there is no reason to expect that the same paradigms, techniques, and approaches that have been so successful in previous robotics applications will also succeed here. Modeling human behavior is a complex and altogether different task, and it is reasonable to expect that new algorithmic approaches will have to be developed to address it. In developing these new approaches, it would behoove us to consider the existing work in related fields such as psychology, even if their theories are not expressed in the form of computational models.

A lot of early HRI focused on physical design and social psychology involving people's perceptions of robots. While those are interesting topics, especially for exploratory research, the overall agenda of HRI research must expand to include building systems that can perform interesting or useful tasks. One good thing about this early focus is that the extreme importance of evaluating systems through their interaction with real humans was immediately established. Simulation makes development easier, but we know that the models of human behavior that we work with are most likely to incomplete for the purposes of rigorous evaluation. People's subjective impressions of an interaction are important, but the ultimate goal of making robots that are compelling and interpretable to people is so that they may better perform in the roles that they are assigned to. With this in mind, it is important that the tasks we address in our research,whether they are utilitarian or social, have objective measures for success or failure that can be evaluated.

Robotics relies on the synthesis of technologies from related fields, and HRI is no different. In order that HRI not be a closed dialog, researchers need to advertise HRI to the greater robotics community as a sub-discipline full of new and interesting application domains. Additionally, it is necessary that we open a dialog between HRI researchers and researchers who develop technologies on which effective HRI systems may rely (vision, dialog management, etc) to express a need for robust, real-time, and in particular, human-centered systems.

My own research focuses on planning for HRI in *socially situated tasks*. Socially situated tasks are defined to have these characteristics:

- Have goals whose attainment are objectively verifiable and rooted in the physical world (ie, "get to the end of the hallway", not "make friends with this person")
- Are performed in the presence of other social actors and require some level of interaction or coordination in order to succeed
- Have familiar protocols for the achievement of goals (social conventions) that are assumed to be known by all participants

Examples of some socially situated tasks are navigating through crowded hallways, riding elevators with other people, and driving in traffic. In these types of interactions, people use their knowledge of social conventions to reason about the intentions of the people they are interacting with. The behavior that they choose to achieve their goals are based both on their knowledge of social convention and their beliefs about the goals of other people.

The planning paradigm used in this research is decision-theoretic planning. The interactions are modeled as partially observable Markov decision processes (POMDPs), in which other peoples' intentions are treated as hidden state. The model itself is developed by combining a structure based on expert knowledge with real data gathered from people performing the interaction. Probabilistic planning has been used with great success in other robotics applications. The flexibility of probabilistic models make them a good choice for representing complex, real-world behaviors. Additionally, the hidden state in these models is a natural way to represent the influence of people's internal mental states, which cannot be directly observed, on their actions.

Additionally, we suspect that because people's actions may depend on the passage of time during an interaction, completely abstracting time away (as is typically done in decision-theoretic planning) may be inadequate to capture essential characteristics of some human interactions. This hypothesis will be tested by comparing the performance of models that explicitly represent time in the model with more traditionally used fully Markov models.

The example interaction domain in which this approach is being tested is the Pittsburgh Left. The Pittsburgh Left is an unofficial, local driving convention wherein a car turning left from a street without a turning lane may be allowed to go before other cars immediately after the light turns green

F. Broz is a graduate student at the Robotics Institute, Carnegie Mellon University, 5000 Forbes Ave Pittsburgh, PA 15213, USA fbroz@cmu.edu

if the oncoming driver decides to allow them. In a first set of experiments, data was collected of people's driving behaviors while taking (or not taking) and allowing (or disallowing) the Pittsburgh Left in a driving simulator game.The data was used to create various POMDP models of the interaction. The relationship between both drivers' intentions and the outcomes of the trials (which car exited the intersection first) show interesting trends in how people resolve conflicting goals in this particular interaction.

The policies produced by solving these models will be evaluated through a second set of driving simulator experiments, in which people will interact with agents controlled by these policies. The method of evaluation will be two-fold. User studies will be employed to examine people's subjective impressions of interacting with the agents and evaluate the perception of social appropriateness of the agents' behavior. The objective criteria of the amount of reward, based on the trial outcome, achieved by the agents and humans during these interactions will also be analyzed. This will allow both the agents' and the humans' reward to be compared to the rewards achieved by humans interacting with other humans during the first set of experiments. It is our belief that it is necessary that an agent be judged as socially competent by the human interacting with it in order to perform well at this task, and that is the hypothesis that this method of analysis seeks to confirm or deny.