

Social Feedback for Robotic Collaboration

Emily Wu, Brown University

October 2, 2015

1 Introduction

Collaboration is a process that relies heavily on communication for its success. When humans collaborate, this communication is both obvious and implicit—we not only instruct and request aid from each other, but we also tacitly monitor our partners for signs of approval and understanding while producing these signals ourselves. However, in robotics, many of these vital components of successful communication are missing or lost. The absence of these signals likely account for numerous failures in human-robot collaborative tasks. I propose to create a framework that allows for these missing components to be added to human-robot collaborative tasks, which should result in faster, more accurate, and less frustrating human-robot collaboration experience.

To describe this process of back-and-forth signaling, we use the term *social feedback*. Social feedback refers to the implicit and explicit communication and monitoring of state. Explicit feedback is often made through questions or statements, e.g., “Do you want this one?” (explicit monitoring) or “Sorry, I didn’t understand” (explicit communication). Implicit feedback are more subtle, such as looking at a particular object or making a confused expression (implicit communication) or merely observing your partner for understanding (implicit monitoring). We aim to arm our robot with these communicative skills.

By enabling robots with social feedback, we hope to see several improvements in many domains. Overall, we expect a better human-robot interactive experience, as the robot will feel more responsive to human feedback by providing cues of understanding, and be more engaging as it actively communicates with the human, instead of simply obeying its commands. Because of this, we also expect novice users to feel more comfortable collaborating with the robot, as the collaboration will be more analogous to that with humans. Additionally, communicating with the human will allow the robot to use the human as a resource to help it recover from failures. Overall, we expect these changes to result in increased speed and accuracy in human-robot tasks.

2 Existing Work

Previous work has been published towards this goal. In our **The paper**, we describe a Partially Observable Markov Decision Process (POMDP) model of an object handoff task where the robot is capable of explicit monitoring—i.e., asking yes or no questions.

2.1 Domain: Object Handoff

The domain we tested our system on involves the robot handing a requested object to a human. Our Baxter robot is situated by a table with objects located in reach. A human participant requests Baxter hand them an object using natural language and gesture. Depending on whether social feedback is enabled, Baxter can ask the human questions before handing the human the desired object.

2.2 POMDPs

We formulate this domain as a POMDP, parameterized by seven variables, $\{S, A, T, R, \Omega, O\}$.

S describes the set of states, which in our domain is a tuple (ω, \mathcal{O}) . The set of objects on the table is \mathcal{O} , and $\omega \in \mathcal{O}$ is the object the human desires. Each object in \mathcal{O} is parameterized by its name, its location, and a vocabulary of words that describe it. The desired object ω is not known, and can only be inferred via observations. However, we consider \mathcal{O} to be completely observable and to change deterministically.

A is the set of actions that our robot can take. This set always includes the actions *PICK*, where the robot picks an object and hands it to the human, and *WAIT*, where the robot does nothing. If social feedback is enabled, the robot is also capable of choosing the action *ASK*(x) where x is a property derived from one of the words of the object vocabulary.

T is the transition function, which describes the probability of transitioning to another state given an action: $p(s'|s, a)$. In this domain, we restrict the set of objects on the table to only change if an object is picked up. Similarly, the desired object only changes if that object is picked, and probability is redistributed uniformly across the remaining objects.

R is the reward function, which takes as input a state and an action and outputs a real-valued reward, incentivizing and disincentivizing the robot appropriately. In this domain, we choose a high positive reward for the robot picking the correct object, and a large negative reward for picking the incorrect object. In addition, we reward *WAIT* actions with a small negative reward and any *ASK*(x) actions are given a slightly higher penalty to disincentivize the robot from asking too many questions and annoying the human.

Ω describes the set of all observations. Observations here are a tuple of (l, g) , where l is the human’s language as interpreted by speech recognition, and

g is the human’s gesture, as determined through the position of the human’s shoulder and wrist compared to the location of objects on the table.

O is the observation function, which gives the probability $p(o|s, a)$ —the probability of witnessing a certain observation given the state that it occurred in and the action that triggered it. We consider language and gesture to be conditionally independent

3 Continuing Work

4 Results

5 Conclusion