Learning from the Wizard: Programming Social Interaction through Teleoperated Demonstrations

ABSTRACT

This paper considers the question of whether robots can be effectively programmed for autonomous social interaction through learning from demonstration. We advocate for learning from demonstrations recorded via Wizard-of-Oz teleoperation. First, we give an overview of this paradigm of learning from the wizard (LfW). Then, we present a novel LfW system for educational play between young children and a robot. We describe initial results from a randomized experiment comparing a teleoperated robot and an autonomous robot whose behavior was derived by LfW. Across numerous metrics, the teleoperated robot and the autonomous robot programmed by LfW elicit similar behavior from their human interaction partners. Additionally, when children were asked whether the robot was human-controlled or autonomous, approximately half in each condition thought it was human-controlled. These results demonstrate that a robot programmed via LfW can successfully engage in highly social, interactive tasks with young children.

Categories and Subject Descriptors

I.2.9 [**Artificial Intelligence**]: Robotics

General Terms

Experimentation,Algorithms

Keywords

socially assistive robots, learning from demonstration, Wizard of Oz, autonomous social interaction

1. INTRODUCTION

In human-computer interaction research, a Wizard-of-Oz scenario involves a human interacting with some machine interface—such as a socially expressive robot—that is secretly controlled by a human teleoperator. We define **learning from the wizard** (LfW) as a subtype of learning from demonstration in which the training demonstrations are derived from records of Wizard-of-Oz interaction and are used to learn policies for interacting autonomously with humans. Previous work on LfW exists but until now has lacked a unifying name and has not been extended to the field of Human-Robot Interaction (HRI); we situate the contributions of this paper amongst past work in Section 2.

This paper focuses on validating LfW by answering *whether learning from the wizard can program effective behavior for social interaction* and *whether that behavior is judged similarly to Wizard-of-Oz behavior by human interaction partners*. To this end, we conducted an evaluation of robot LfW in an educational domain with young children.

2. RELATED WORK

Much of the research involving learning socially interactive behavior from demonstration has been conducted with software agents, particularly with dialog agents. Learning from the wizard has been used to develop dialog systems that act as tour guides [**?**], provide directions to drivers [**?**], and decide whether to provide information on a screen or verbally in an in-car music player [**?**]. This pioneering work contains a wide range of methods and evaluations. None of the work, however, reports quantitative results with the learned agent acting with full autonomy. For example, in some cases, evaluations focused on prediction accuracy [**?**, **?**] on Wizard-of-Oz data, which is informative but not a substitute for testing the learned system in its own interaction, where it may lead to new situations that are seldom encountered during Wizard-of-Oz control.

3. LfW FOR AN EDUCATIONAL DOMAIN

To explore the potential of LfW for social tasks, we developed a robotic learning companion for young children. This section describes the educational domain and integrated system.

Figure 1 shows the interaction setting. Children aged 4–8 sat across from a squash-and-stretch robot that resembles a small plush toy and was designed for social interaction with young children. Between the robot and child was a tablet with an educational app (also shown in Figure 1) designed to teach color-mixing and reading concepts to young children. Children could tap on leaves to add color to the birdbath, where it mixes with previously added colors, and to wash the bath clean, and they could touch the words to hear them read aloud. Adding multiple pigments to the birdbath can create secondary colors and brown; a video of app interaction can be viewed at `https://www.youtube.com/user/LearningFromWizard`.

3.1 General system

Appears in: *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016), John Thangarajah, Karl Tuyls, Stacy Marsella, Catholijn Jonker (eds.), May 9–13, 2016, Singapore.*

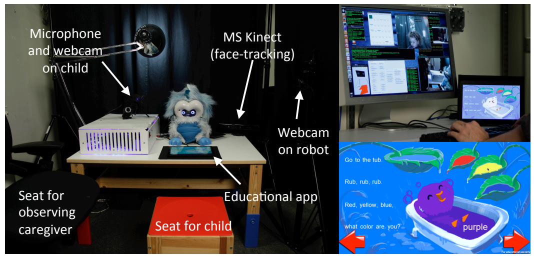

Figure 1: Clockwise from left: Interaction setting; experimenter interface; color-mixing app, from the Tinkrbook e-book.

The system sensors include 2 webcams (one pointed at the child, another at the robot), a Microsoft Kinect, a microphone, and the tablet itself. The tablet reports the start and end of child touches by their locations and times. It also reports time-stamped app events, such as triggering pigment to fall or a word of text to be said aloud.

The final system, used during the experiment, has 30 robot actions. 15 are simple emotional utterances paired with expressive motions such as an excited "Yeah!" paired with an quick up-and-down motion, an impatient "Ooh ooh!", or a whimper. 5 actions cause the robot to lean towards the tablet and trigger app events. These app events are: adding 1 of 3 primary colors, washing the birdbath clean, or tapping the tub. 9 actions are prompts to the child. 8 of these ask for a possible color in an excited and anticipatory tone (e.g., "Brown?!", "Red?!", or "White?!"), and the 9th is requesting "Now you go." A recording of the robot performing each action can be viewed at `https://www.youtube.com/user/LearningFromWizard`.

During each interaction, the various system components communicated via ROS. Time-stamped logs in the form of rosbags were recorded that contained video from the two webcams, app events (human- or robot-triggered), estimates of the child's face location in 3D space (derived from the Kinect), microphone audio, and robot actions.

3.2 Autonomy by learning from the wizard

In this section, we give a brief overview of how we structured teleoperation, training, and learned autonomous behavior. The robot's actions are exclusively determined either by a teleoperator during Wizard-of-Oz interaction or by a learned policy during autonomous interaction, never by a mixture of the two.

To create the autonomous behavior model, a set of features (e.g., the current color of the birdbath, the child's and robot's last actions, time since last child/robot action, etc.) are computed from logs of teleoperated data during learning .

Our goal is to learn a model that answers: given a feature vector $\boldsymbol{f}$ drawn from sensory history at time t, for each action a, what is the probability $\pi(\boldsymbol{f}, a)$ that a would have been triggered by a teleoperator within the previous 100ms? For this 30-class learning problem, we created a hierarchical model using binary logistic regression (from Weka) as a base learner, with L2 regularization applied with a scaling parameter of 10 to account for the large number of features.

During autonomous interaction, these same features are computed at the start of each 100 ms time step, and a single action is chosen and executed by sampling from the model's output distribution over actions, $\pi(\boldsymbol{f}, \cdot)$.

4. STUDY DESIGN

Data from 37 participants' interactions with a teleoperated robot over a 4-day period formed the training data. A separate group of 85 participants took part in the final stage, a randomized experiment with three conditions, differing in what the child played with: TABLET-ONLY, a control condition without the robot; WOZ-EXPERIMENT, with the teleoperated robot; and AUTONOMY, with the robot acting according to its learned autonomous model.

5. RESULTS AND DISCUSSION

The main research question of this paper is whether LfW can be used to program successful socially interactive behavior. One way to answer that question is to assess how well the learned behavior emulates the successful social behavior of the teleoperator. For this assessment, we focus on key behavioral metrics of the child's interaction with either a teleoperated robot or an autonomous robot acting according to the learned policy during a randomized experiment between learned autonomy and Wizard-of-Oz control.

5.1 Learned behavior

In the authors' subjective judgement, the robot largely succeeded in learning to emulate the demonstrated interaction heuristics. For instance, the autonomous robot often celebrates when the child creates a robot-requested color. Videos of learned and Wizard-of-Oz behavior can be viewed at `https://www.youtube.com/user/LearningFromWizard`. We describe objective, quantitative results in the following section.

5.2 Participants guess: teleoperated or autonomous

As a double-blind question, another experimenter asked, "When you were playing with [the robot], [the experimenter] was sitting at the computer behind you. Do you think he was remote-controlling [the robot], telling him exactly what to do, or do you think he was just starting and stopping the robot?" I.e., they were asked whether the robot was autonomous or teleoperated by a human. Through this question, the experiment serves as a variant of the Turing test [**?**], albeit social, embodied, and constrained to 30 possible actions. Counting answers of "undecided" or "both" as half a vote for each option, 52.8% of respondents in the AUTONOMY condition thought the robot was teleoperated. 47.2% in the WOZ-EXPERIMENT condition thought it was teleoperated. In other words, the autonomous robot was judged as human-teleoperated marginally more often than the robot that was actually human-teleoperated. A 2x2 Fisher's exact test results in $p = 1$.

Both behavioral metrics and the children's guesses about whether the robot was autonomous indicate that LfW did program social behavior that largely resembled Wizard-of-Oz teleoperation by the demonstrator and was successful compared to child-tablet interaction without a robot.

6. CONCLUSION

From the perspective of *how to program autonomous socially interactive robot behavior*, learning from the wizard constitutes a nascent approach that could be instrumental in the much-needed transition of social robots from Wizard-of-Oz control towards autonomy. Evaluating LfW within a playful, educational domain, this paper provides the first validation of the technique to program socially interactive robot behavior.