Learning from the Wizard: Programming Social Interaction through Teleoperated Demonstrations
EXTENDED ABSTRACT

W. Bradley Knox, Samuel Spaulding, and Cynthia Breazeal
Massachusetts Institute of Technology
wbradleyknox@gmail.com, samuelsp@media.mit.edu, and cynthiab@media.mit.edu

ABSTRACT

This paper considers the question of whether robots can be effectively programmed for autonomous social interaction through learning from demonstrations recorded via Wizard-of-Oz teleoperation. We present a novel LfW system for educational play between young children and a robot and results from a randomized experiment comparing a teleoperated robot and a robot with autonomous behavior 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.

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 learn tasks of social human-robot interaction.

This paper focuses on validating LfW by answering *whether learning from the wizard can retain the interaction benefits of a teleoporated robot*. To this end, we conducted an evaluation of robot LfW in an educational domain with young children aged 4–8 years.

2. RELATED WORK

There exist related LfW projects that involve human-interaction tasks, social tasks, or rigorous evaluations of learned behavior [5, 10, 7, 8, 11, 4, 3], but to our knowledge only Breazeal et al.'s work [1] involves all three *on a robot*. However, though their nearly autonomous robot could effectively complete simple, single-agent tasks, its social behaviors—communicating and coordinating with the human—were considered unsuccessful by the authors, leaving open the question of whether LfW can effectively program socially interactive robot behavior.

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

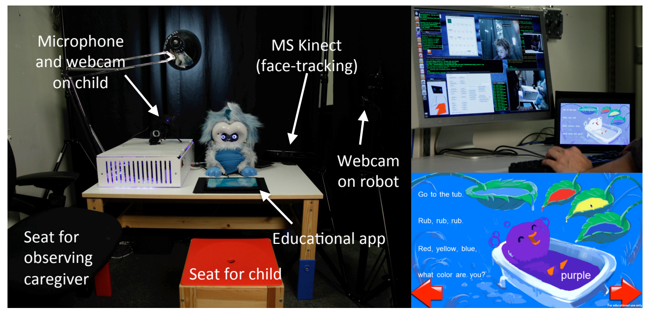

Figure 1: Clockwise from left: Interaction setting; experimenter interface; color-mixing app, from the Tinkrbook e-book.

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 was designed for social interaction with young children. Between the robot and child was a tablet with an educational app designed to teach color-mixing and reading concepts to young children. Children could tap on leaves to add color to the birdbath and to wash the bath clean. Adding multiple pigments to the birdbath can create secondary colors and brown.

3.1 General system

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. During each interaction, the various system components communicated via ROS [6]. 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.

The final system, used during the experiment, has 30 robot actions in its agent framework. 15 are simple emotional utterances paired with expressive motions. 5 actions cause the robot to lean towards the tablet and trigger app events. 9 actions are prompts to the child, requesting one of eight colors (e.g., "Red?!") or "Now you go." The final action is to not initiate any behavior. Video illustrations of app events, the robot action space, and child-robot interaction can be seen at `https://www.youtube.com/user/LearningFromWizard`.

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, 122 feature values (e.g., describing the current color of the birdbath, the child's and robot's last actions, time since last child/robot action, etc.) are computed at 100ms time intervals from logs of teleoperated data. Each set of feature values creates a sample for supervised learning, labeled by what action the teleoperator specified over the previous 100ms. 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 [2]) 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 probability distribution over actions, $\pi(\boldsymbol{f}, \cdot)$.

4. STUDY DESIGN

Data from 29 participants' interactions with a teleoperated robot formed the training data. A separate group of 85 participants took part in a subsequent 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

Subjective observations and quantitative analysis are presented below.

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. The following section gives objective, quantitative results.

5.2 Behavioral results

Multiple-testing analysis of 11 behavioral metrics—by the Benjamini and Hochberg procedure, with a false discovery rate of 0.05—results in numerous significant differences between TABLET-ONLY and either AUTONOMY or WOZ-EXPERIMENT and *zero* significant differences between the two robot conditions. Significant results include children that interacted with the robots volunteering to play longer, having a higher proportion of color taps result in creating complex colors, smiling more, and looking down at the screen less. Together, these results indicate that the learned autonomous robot preserves the interaction benefits of the teleoperated robot.

5.3 Participants guess: teleoperated or autonomous

As a double-blind question, another experimenter asked participants whether the robot was autonomous or teleoperated by a human. Through this question, the experiment serves as a variant of the Turing test [9], 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; the autonomous robot was judged as human-teleoperated marginally more often than the robot that *was* human-teleoperated.

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, showing that LfW generally retained the statistically significant interaction benefits that a teleoperated robot has over play without the robot.

REFERENCES

[1] C. Breazeal, N. DePalma, J. Orkin, S. Chernova, and M. Jung. Crowdsourcing human-robot interaction: New methods and system evaluation in a public environment. *Journal of Human-Robot Interaction*, 2(1):82–111, 2013.

[2] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. The weka data mining software: An update. *SIGKDD Explorations*, 11, 2009.

[3] C.-M. Huang and B. Mutlu. Learning-based modeling of multimodal behaviors for humanlike robots. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, pages 57–64. ACM, 2014.

[4] S. Nikolaidis and J. Shah. Human-robot cross-training: Computational formulation, modeling and evaluation of a human team training strategy. In *Proceedings of the IEEE/ACM International Conference on Human-Robot Interaction*, 2013.

[5] M. Okamoto, Y. Yang, and T. Ishida. Wizard of oz method for learning dialog agents. In *Cooperative Information Agents V*, pages 20–25. Springer, 2001.

[6] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng. Ros: an open-source robot operating system. In *ICRA workshop on open source software*, volume 3, page 5, 2009.

[7] V. Rieser and O. Lemon. Learning human multimodal dialogue strategies. *Natural Language Engineering*, 16(01):3–23, 2010.

[8] J. Schrum, I. V. Karpov, and R. Miikkulainen. UtâĄĄ 2: Human-like behavior via neuroevolution of combat behavior and replay of human traces. In *Computational Intelligence and Games (CIG), 2011 IEEE Conference on*, pages 329–336. IEEE, 2011.

[9] A. M. Turing. Computing machinery and intelligence. *Mind*, pages 433–460, 1950.

[10] J. Williams and S. Young. Using wizard-of-oz simulations to bootstrap reinforcement-learning-based dialog management systems. In *Proceedings of the 4th SIGDIAL Workshop on Discourse and Dialogue*, 2003.

[11] J. Young, K. Ishii, T. Igarashi, and E. Sharlin. Style by demonstration: teaching interactive movement style to robots. In *Proceedings of the 2012 ACM international conference on Intelligent User Interfaces*, pages 41–50. ACM, 2012.