Investigating Machine-Learning Interaction with Wizard-of-Oz Experiments

Rob Sheline, Christopher J. MacLellan
{robert.sheline, chris.maclellan}@soartech.com

Problem

ML research has commonly focused on reducing model loss, but not how natural and efficient these systems are for developers and end users. However, sampling from the space of interaction designs is expensive because each point requires a different underlying intelligent behavior to exist. We believe this represents a bottleneck for building ML systems using a typical iterative-design process.

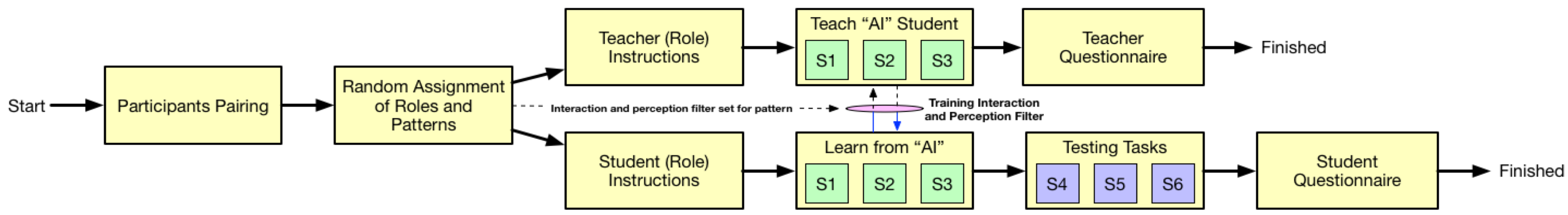

Figure 1: Dual-sided, restricted-perception Wizard-of-Oz experimental design

Our Approach

To alleviate this bottleneck, we extend the notion of *Wizard-of-Oz* (WOZ) experiments from HCI literature. In traditional WOZ experiments, an AI system is simulated by an experimental confederate. We extend this approach to support ML system design by instead replacing the hypothetical learning system with a naïve experimental participant that must learn the task. To prevent the human learner from using out-of-bounds skills and senses, we restrict their perception and available actions to align with the interaction surface that would be available to the hypothetical ML system. This approach, which we call *dual-sided*, *restricted-perception* WOZ, relies on the assumption that the human agent serves as a "good enough" black-box general learner.

Table 1: The Natural Training Interaction Framework

Knowledge	Patterns	Types	Modalities
Goals	Passive Learning	Command	Command Line
Beliefs	Operant Conditioning	Clarify	Control Device
Concepts	Direct Instruction	Acknowledge	GUI
Experiences	Apprentice Learning	Inform	Sketch
Skills	After-Action Review	Spotlight	API
Dispositions	Collaborative Learning	Annotate	Gesture
	Programming	Reward	Speech
		Demonstrate	Text
		Direct Knowledge Manipulation	Multi-Modal
		Request <type>	

We developed this approach to supports the goal of building a general theory of ML interaction design—supporting the design of ML systems that are natural and efficient for humans to teach. Towards this goal, we leverage the *Natural Training Interactions* (NTI) framework (MacLellan et al. 2018), which decomposes the space of ML system interaction designs along four dimensions: knowledge, patterns, types, and modalities (see Table 1).

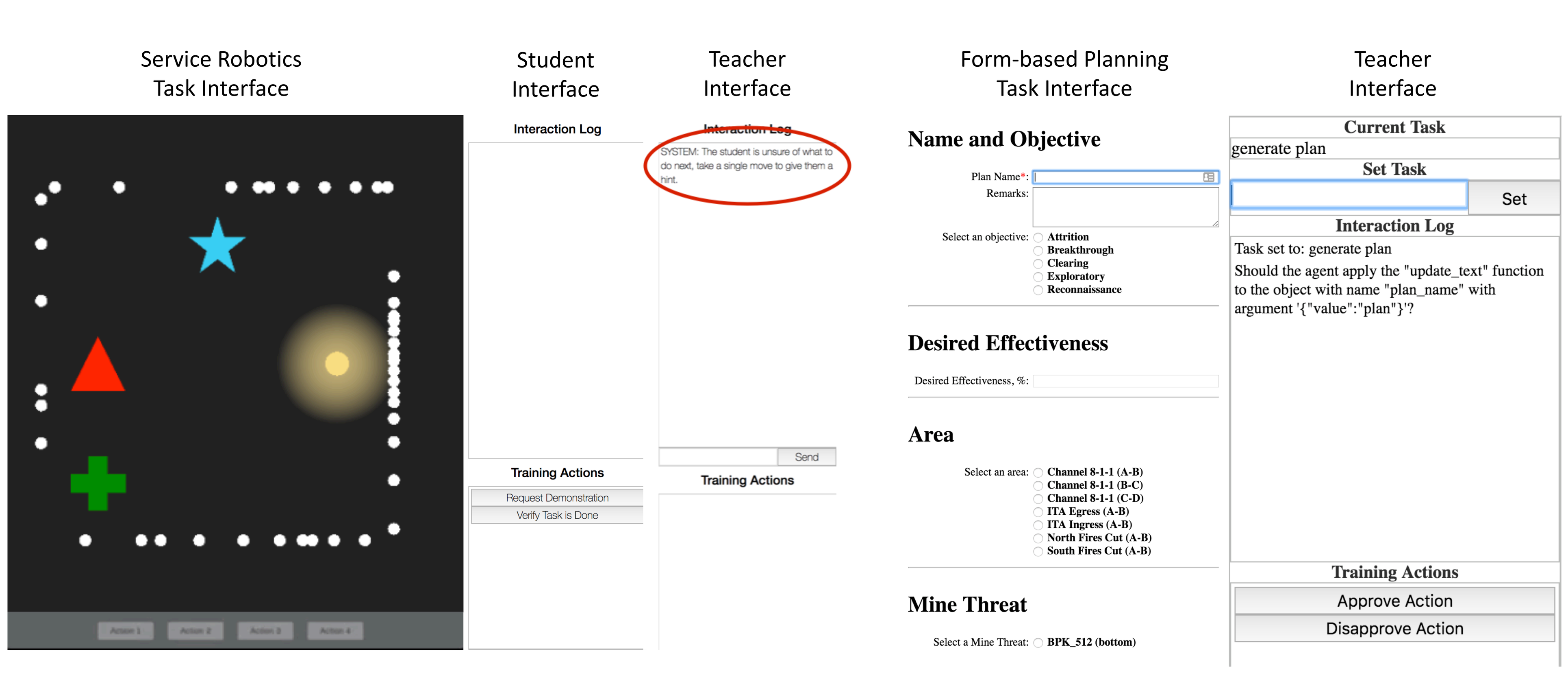

Figure 2: Testbed Task and Teacher/Student Interfaces

Method and Results

To apply this approach, we implemented a web-based testbed for administering WoZ experiments. The goal of the testbed is to enable rapid, cost-effective prototyping of ML interaction designs for new tasks—providing insight into where to direct future ML development efforts. Our testbed implements patterns, types, and modalities in a generic way and provides a generic task API, so it easily supports new interaction designs and tasks. The testbed provides standard UIs for interacting with teachers and students as well as a set of questionnaires (see Table 2) for evaluating the interaction designs.

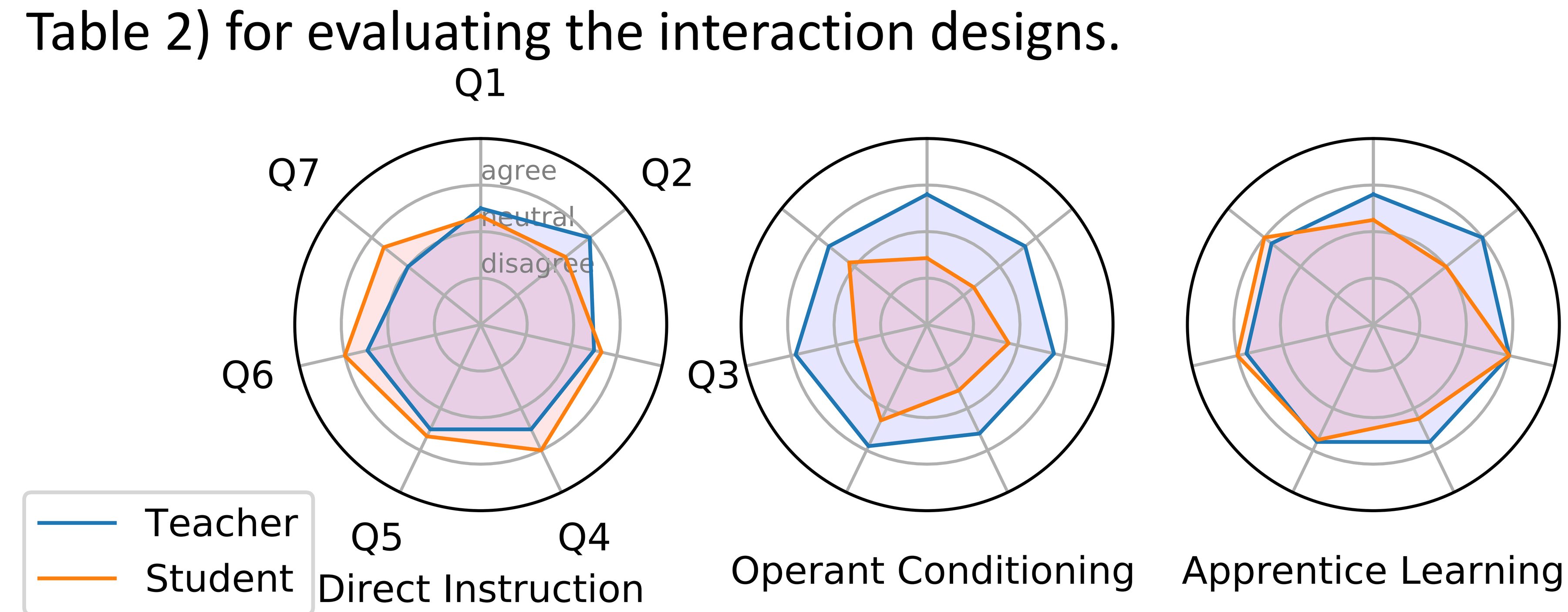

Figure 3: Pilot participant questionnaire responses for service robotics task.

To test the experimental framework and protocol, we created a simple service robotics task (see Figure 2), where agents can explore a maze-like environment, pick up, and put down objects. Next, we designed three hypothetical ML systems that use three patterns from the NTI framework (operant conditioning, direct instruction, and apprentice learning). Figure 3 depicts aggregated answers to exit survey questions Q1-Q7 (from Table 2). This preliminary work serves as a proof of concept that WoZ experiments can be applied to study machine learning interaction designs.

Table 2: The Teacher and Student Questionnaires

Question	Student	Teacher
Q1	I learned to correctly perform the task by the end of training.	The AI student learned to correctly perform the task by the end of training.
Q2	I only needed a few examples to learn.	The AI student only needed a few examples to learn.
Q3	I was able to quickly decide what actions to take next.	The AI student quickly decided what actions to take next.
Q4	Learning from the AI teacher was natural and intuitive.	Instructing to the AI student was natural and intuitive.
Q5	The instructional feedback provided by the AI teacher was always useful.	The instructional options I was presented with were always useful.
Q6	I always received the instruction from the AI teacher that I wanted.	I was always provided with all the instructional options I wanted.
Q7	Learning from the AI teacher was easy.	Instructing the AI student was easy.

Acknowledgements

This material is based upon work supported by the OFFICE OF NAVAL RESEARCH under Contract No. N68335-18-C-0401. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the OFFICE OF NAVAL RESEARCH.

SOARTECH