Towards Understanding How Humans Teach Robots

Tasneem Kaochar, Raquel Torres Peralta, Clayton T. Morrison, Ian R. Fasel, Thomas J. Walsh, and Paul R. Cohen

The University of Arizona, Department of Computer Science,
Gould-Simpson Building, 1040 E. 4th Street, Tucson, AZ 85721-0077
{tkaochar,rtorres,clayton,ianfasel,twalsh,cohen}@cs.arizona.edu

Abstract. Our goal is to develop methods for non-experts to teach complex behaviors to autonomous agents (such as robots) by accommodating "natural" forms of human teaching. We built a prototype interface allowing humans to teach a simulated robot a complex task using several techniques and report the results of 44 human participants using this interface. We found that teaching styles varied considerably but can be roughly categorized based on the types of interaction, patterns of testing, and general organization of the lessons given by the teacher. Our study contributes to a better understanding of human teaching patterns and makes specific recommendations for future human-robot interaction systems.

1 Introduction

Robots and other intelligent devices capable of carrying out highly complex procedures are becoming ubiquitous in the home and workplace. However, changing the behavior and capabilities of these devices typically requires direct programming by specially trained engineers. While machine learning (ML) algorithms offer the allure of allowing machines to improve their knowledge and behavior from experience, ML algorithms still require considerable expertise to use in practice.

To bridge this gap, *human-instructable computing* seeks to develop intelligent devices that can be *taught* by natural human instruction. By "natural," we mean those patterns of communication that humans use every day while teaching each other. For example, humans provide explicit *definitions* as well as *examples* of `concepts`, *describe* and provide *demonstrations* of `procedures`, give *examples* or *definitions* of `rules` and `conditions`, and provide various kinds of *feedback* based on student behavior. We refer to these patterns of instruction as *natural instruction* methods or *natural modes of interaction*. The focus of this paper is on understanding what it would take to automatically map natural human instructions to state-of-the-art ML techniques, so that we can incorporate ML into an end-to-end human-instructable machine.

Prior work in machine learning has studied aspects of human-instructable computing through the lens of single instruction modes such as demonstration [1], teaching concepts by examples [4,5], and human-provided reinforcement [3,6]. In each of these, instruction sessions must be carefully set-up by an expert. In this paper we tested a new interface that allows the user to flexibly switch between all 3 natural instruction modes as well as an explicit testing phase. This was done with the goal of answering several questions about how a human *teacher* interacts with an *electronic student*: (1) What

Joseph A. Konstan et al. (Eds.): UMAP 2011, LNCS 6787, pp. 347–352, 2011.
© Springer-Verlag Berlin Heidelberg 2011

natural instruction methods do humans actually use and in what proportion? (2) When and how often do humans switch between instruction modes? (3) How much teaching is implicit rather than explicit? (4) Are there identifiable teaching patterns that we can utilize to better model and design human-robot interaction systems?

Using our new multi-modal instruction interface, we conducted a study of non-expert users and observed how they taught a task in a simulated flight environment to an electronic student they believed was capable of learning. We report on a number of characteristic teaching patterns and styles that were revealed in the teaching sessions.

2 Methodology and Protocol

To answer the above questions, we need to ask novice human participants to teach a series of inter-dependent concepts and tasks to an electronic student. Ideally we would allow humans to use any form of interaction they like while a system learned from their instruction in real-time. However, the current sophistication of natural language processing is insufficient for this task and general ML systems that learn from natural human interaction have not yet been developed.

We therefore designed an interface that enables subjects to flexibly choose among a variety of teaching methods, but using interface elements that can plausibly be interpreted by state-of-the-art machine learning algorithms without requiring natural language understanding. We then asked 44 University of Arizona students to interact and teach an "electronic student" using our interface; to simulate a competent learning agent, the electronic student was actually secretly controlled by a confederate human (the *wizard*), in a so-called "Wizard of OZ" protocol.

Wizard of OZ Experimental Setup: Each teaching session consisted of a participant, who had no prior knowledge of the goals of the project, taking on the role of the *Teacher* while a researcher played the role of the *Student*. The Teacher was led to believe that he/she was interacting with an electronic student. The Teacher was first trained (by a second *Experimenter*) on the use of the interface and then presented with the teaching task outlining the knowledge the Student should attain by the end of the teaching session. The actual teaching sessions lasted from 25 to 35 minutes. The same two researchers took on the role of the Student and the Experimenter, respectively, across all experiments in order to ensure consistent training and Student/Wizard behavior.

The task took place in an Intelligence, Surveillance and Reconnaissance (ISR) domain in which the Student controls a simulated unmanned aerial vehicle (UAV) and is taught to carry out missions. The simulated environment includes a terrain map with objects that can be scanned using two sensors: a *high-resolution camera* (provides detailed object information, such as whether a boat has a cargo hold), and a *radiation sensor* (detects the radiation level of an object in range). The Student can only perceive the world through the sensors and the Teacher must teach the Student how to use the UAV sensors in the appropriate circumstances.

Our interface provides the Teacher three tools to teach the Student: (A) the Instruction Command Interface (ICI), which sends commands to the Student; (B) a Timeline Display that shows a list of all prior Teacher instructions; and (C) a Map Display providing information about world objects, UAV sensor state and range and UAV flight path. Action commands (from the ICI) direct the Student's control of the UAV and examples include *use camera to track object*, *fly to location*, etc. Four modes of instruction are supported by the ICI: (1) *Teaching by demonstration*: Teacher can label a sequence of actions as an example of a procedure. This can be done either by labeling the beginning and end of the procedure while providing a demonstration, or after the fact by selecting already executed actions from the timeline. Teacher can provide multiple examples of a single procedure and each instance can be labeled as a positive or negative trace of a specific procedure demonstration. (2) *Teaching concepts by examples*: Teacher can define object concepts (such as "cargo boat") by selecting an object on the map interface and giving it a label. Again, positive and negative examples of an object label can be given. (3) *Teaching by reinforcement*: Teacher can give feedback to Student at any time, in the form of 1-3 "happy faces" or 1-3 "frowny faces". Teacher can also label goals and indicate when they have been met. (4) *Testing*: Teacher can test the Student's learning by giving commands that ask the Student to provide a label for an object or execute a previously defined procedure.

Teaching Task: In each of the teaching sessions, there were two kinds of objects, cargo boats and fishing boats; the Teacher was asked to teach the Student how to distinguish them, to use the radiation sensor only on cargo boats, and to generate a report of the readings. Although this task is very simple, it requires teaching multiple object concepts and procedures that depend on one another. Teaching sessions were recorded and a transcript of the Teacher-Student interaction was generated for each teaching session.

3 Results: Analysis of Transcripts

We analyzed the transcripts with the aim of answering the key questions raised in Section 1. We discovered a number of quantifiable patterns, including three major findings:

1. Humans use multiple modes of instruction and these modes are often tightly interleaved. (This observation appears to be independent of the teaching task since it was noted also in our prior pilot studies [2]).
2. We found at least 4 distinct patterns used to switch between teaching and testing of the electronic student.
3. We observed at least 3 categories of human teaching "style" based on the level of organization of the Teacher's instructions.

3.1 Modes of Instruction

In order to answer our first question regarding how often natural instruction methods are used by human teachers, we counted the occurrences of the *label object* (mode: teaching concept by example), *define procedure* (mode: teaching by demonstration) and *give feedback* (mode: teaching by reinforcement) constructs. We found that more than half of our participants (57%) made use of all three modes of teaching while 32% taught using only two modes of instruction (7% demonstration and concepts, 14% demonstration and reinforcement, 11% concepts and reinforcement). 11% taught using only teaching by demonstration (2%) or concepts by examples (9%), the latter of which was insufficient for completing the teaching task. Teaching by reinforcement was never employed by itself and in the 82% of cases where it was used, it followed another teaching type in all but 2 cases. Fifty-eight percent of the teachers who used feedback used it exclusively after testing. This indicates that reinforcement feedback is most useful in this task for fine tuning behavior that has been "bootstrapped" with other instruction modes. When teaching concepts by examples, 75% of participants used positive *and* negative examples, while 25% only used positive examples, which were sufficient for the teaching task. This evidence suggests that humans tend to provide exceptions along with rules.

In contrast to the relatively simple object concept teaching, our studies showed that teaching a procedure is non-trivial. For example, when teaching procedure definitions, we found that teachers do not always declare the procedure up front. While 60% of participants declared the procedure before beginning to teach it, 20% of the teachers identified the procedure only after providing a demonstration. The remaining 20% of teachers vacillated between both styles. Furthermore, 41% of our teachers never explicitly defined a procedure, even though teaching at least one procedure was required to complete the teaching task. In our post-study questionnaire, 41% identified the object labeling construct as "easy to use", compared to only 16% for the procedure construct, and 23% identified teaching a procedure as a difficult task. Finally, we observed that in most transcripts, some procedures were taught implicitly. In these cases, the Teacher directed the Student to perform the same sequence of actions repeatedly in different locations of the world but never explicitly declared that a procedure was being taught.

We also split each teaching session into 3 equal time phases and analyzed whether the usage of instruction modes changed over time. Even though a teaching or testing instruction may continue across several phases (such as a procedure demonstration), it was only classified under the time phase in which it was *started*. We found that teaching concepts by example was prevalent throughout the teaching session (61%, 45%, 59%), and 84% of the teachers *began* their session by labeling objects. In contrast, teaching by demonstration (43%, 41%, 64%) and by reinforcement feedback (27%, 36%, 68%) increased in the later phases. We believe this pattern demonstrates a "bootstrapping" technique as most teachers attempted to teach procedures later in the teaching session based on the object labels taught earlier. Moreover, the steady increase of reinforcement feedback over time reflects the effects of testing the Student and providing feedback to fine-tune its behavior in the later phases.

3.2 Teaching and Testing Patterns

Teachers testing their students is an important facet of Teacher-Student interaction. All but 6 of our participants made use of the testing tools at some point during their teaching sessions. Using the 3 temporal phases described above, we observed that while teaching tools were used throughout the teaching session (84%, 68%, 93%), testing tools were most popular in the third phase (41%, 43%, 75%). We also found 4 distinct patterns of teaching and testing employed by our participants. **Type A** teachers always test the Student's comprehension after teaching a concept or procedure and before teaching another new concept. **Type B** teachers loosely interleave teaching and testing, introducing several new concepts to the Student at a time before doing any testing. **Type C** teachers reserve all testing for the end of the session whereas **Type D** teachers do not test at all.

We found that half of our participants fell under Type A, followed by Type B (25%), Type D (14%), and finally, Type C (11%). We hypothesize that Type A teaching is an indication that a teacher is uncertain about how the Student actually learns. We noticed 2 teachers who began the session with Type B but then switched to Type A after the Student failed a test. Furthermore, 4 teachers tested the Student *before* doing any teaching (perhaps in an attempt to understand the Student's base knowledge). Finally, it was common for teachers to give feedback during or immediately following a test protocol, either to express satisfaction or disappointment, or to complement the teaching.

3.3 Teaching Styles

Our transcript analysis also revealed distinct styles of teaching based on the *organization* of lessons. *Structured* teachers (16% of participants) were consistent and methodical in the execution of their instruction commands. They consistently used the interface's object labeling construct to teach object concepts and the procedure demonstration construct to define procedures. These teachers always tested Student's comprehension after teaching a lesson. *Semi-structured* teachers (50%) began with a less structured teaching style but became progressively more structured as the teaching session continued. They made use of the GUI features almost as intended, sometimes with early exploration of usage. *Free style* teachers (34%) were the most difficult to follow, mainly because these teachers made use of GUI features in novel ways. A few of these teachers tested Student's knowledge of world object labels *before* doing any teaching. Four teachers even appeared to use the procedure testing tool to provide further teaching examples of a procedure originally taught via the procedure definition construct.

One novel use of the procedure construct was to use it as a concept labeling device. That is, free-style teachers might define two separate procedures with the same sequence of actions, yet give them different names ("cargo boat" and "fishing boat") in an apparent attempt to teach the labeling distinction through procedure names (25% of participants did so). This stands as a warning to interface designers who might try to tailor an aspect of the GUI to a single mode of instruction – users may find new ways to use UI elements.

Another unexpected usage of the interface was the use of deictic or pointing actions to teach. This often happened in concert with the labeling of concept examples. While structured teachers would fly the plane to an object and then label it using the concept labeling tool (a process that expressly involved clicking on the object), we saw freestyle teachers often using procedure names or other unintended methods to label a concept that was in the *vicinity* of the UAV. The lesson here is that interface designers should be aware that human teachers may expect a certain amount of spatial reasoning performed by the interface or the electronic student on the other side.

4 Conclusions and Discussion

To the best of our knowledge, BLUI's teaching interface is the first to simultaneously support several modes of Teacher-Student interaction over the agent's lifetime: teaching concepts by example, through demonstration and via reinforcement, and testing Student's learning. Looking back at the questions we posed in Section 1, we found that in over half of the teaching sessions, all three modes of instruction were used and the switch between instruction modes usually indicated a "bootstrapping" interplay, such as the use of reinforcement to fine tune behavior previously taught using one of the other instruction modes. Our data suggests that teachers view testing as a a critical part of teaching; we hypothesize that testing helped assure teachers that Student understood what was being taught. Teachers preferred to test the Student intermittently throughout the teaching session rather than doing a monolithic testing episode at the end. The importance of testing in teaching was also observed in our pilot studies, using different teaching tasks [2]. We catalogued several levels of organization that characterized teaching trajectories, and noted that teachers frequently used the GUI in unexpected ways. Finally, while we found that much of the human teaching using BLUI's instruction interface is explicit, the presence of implicit procedure definitions highlights a challenge for ML algorithms, which typically need carefully aligned instructions.

Based on our observations of teaching patterns, we suggest that teaching interfaces for human-robot interaction should (1) allow for fine-grained testing of student's learning, (2) facilitate a bootstrapped teaching style in which concepts or procedures can be taught with one mode and refined with another (such as feedback), (3) allow teachers to provide positive *and* negative teaching examples, and (4) accommodate teachers who may not declare instructional intent in advance.

Our ultimate goal is to build an electronic student that can learn from natural human instruction. This study sought to illuminate how human teachers behave when interacting with intelligent computer systems or robots. The next step is to develop a system which can parse Teacher-Student interactions automatically, identify the boundaries of lessons, feed them to concept (e.g. ILP [4]) or procedure (e.g. planning operator [7]) learners, and then improve upon those learned concepts when feedback is given. However, extracting each teaching episode automatically (without human facilitation) is a non-trivial task, as exemplified by the evidence presented in this study of the different types of teaching-testing patterns and teaching styles naturally used by human teachers.

Acknowledgments. This work was supported by DARPA under contract 27001328 and ONR "Science of Autonomy" under contract N00014-09-1-0658.

References

1. Argall, B., Chernova, S., Veloso, M.M., Browning, B.: A survey of robot learning from demonstration. Robotics and Autonomous Systems 57(5), 469–483 (2009)
2. Kaochar, T., Peralta, R.T., Morrison, C.T., Walsh, T.J., Fasel, I.R., Beyon, S., Tran, A., Wright, J., Cohen, P.R.: Human natural instruction of a simulated electronic student. In: AAAI Spring Symposium (2011)
3. Knox, W.B., Stone, P.: Combining manual feedback with subsequent mdp reward signals for reinforcement learning. In: AAMAS (2010)
4. Natarajan, S., Kunapuli, G., Maclin, R., Page, D., O'Reilly, C., Walker, T., Shavlik, J.: Learning from human teachers: Issues and challenges for ILP in bootstrap learning. In: AAMAS Workshop on Agents Learning Interactively from Human Teachers (2010)
5. Stumpf, S., Rajaram, V., Li, L., Wong, W.-K., Burnett, M., Dietterich, T., Sullivan, E., Herlocker, J.: Interacting meaningfully with machine learning systems: Three experiments. International Journal of Human-Computer Studies 67(8), 639–662 (2009)
6. Thomaz, A.L., Breazeal, C.: Teachable robots: Understanding human teaching behavior to build more effective robot learners. Artificial Intelligence 172(6-7), 716–737 (2008)
7. Walsh, T.J., Subramanian, K., Littman, M.L., Diuk, C.: Generalizing apprenticeship learning across hypothesis classes. In: ICML (2010)