Designing Interactions for Robot Active Learners

Maya Cakmak, Crystal Chao, and Andrea L. Thomaz

***Abstract*—This paper addresses some of the problems that arise when applying active learning to the context of human–robot interaction (HRI). Active learning is an attractive strategy for robot learners because it has the potential to improve the accuracy and the speed of learning, but it can cause issues from an interaction perspective. Here we present three interaction modes that enable a robot to use active learning queries. The three modes differ in when they make queries: the first makes a query every turn, the second makes a query only under certain conditions, and the third makes a query only when explicitly requested by the teacher. We conduct an experiment in which 24 human subjects teach concepts to our upper-torso humanoid robot, Simon, in each interaction mode, and we compare these modes against a baseline mode using only passive supervised learning. We report results from both a learning and an interaction perspective. The data show that the three modes using active learning are preferable to the mode using passive supervised learning both in terms of performance and human subject preference, but each mode has advantages and disadvantages. Based on our results, we lay out several guidelines that can inform the design of future robotic systems that use active learning in an HRI setting.**

***Index Terms*—Active learning, human–robot interaction.**

I. Introduction

OUR research targets social robots situated in dynamic human environments such as general assistants in homes, schools, and hospitals. In these scenarios, it would be difficult for a designer to preprogram every possible task that the robot could perform. Our research is about developing ways for these robots to learn the necessary tasks and skills from end users—socially guided machine learning (SG-ML). Because we cannot always expect these users to have extensive experience with machine learning or robotics, we need to design algorithms and systems that take advantage of the ways that they naturally approach the task of teaching.

Passive and active supervised learning are two machine learning paradigms that can enable a robot to learn from a human. In passive supervised learning, the teacher chooses and labels all of the examples for the learner. In the active learning setting, the learner can select examples to learn from through *queries* to the teacher for labels. In both cases, the human teacher serves as the oracle for the robot learner by providing labels for example data. However, the approaches have vastly different implications in a human–robot interaction (HRI) setting. In particular, the queries of an active learner embed the robot in a tightly coupled dyadic interaction.

In this work, we look at the kind of input that human partners are capable of providing and the kind of feedback that they enjoy receiving. Our focus is on designing systems for users who are nonexperts in machine learning and robotics. We examine the teacher's perception of the learning process in order to better understand how to negotiate the balance between learning accuracy, learning speed, and interaction smoothness. This work provides a foundation for developing active learning systems that can be successfully deployed on robots in everyday human environments.

In a previous pilot study [1], we examined the relationship between active learning and transparency. Transparency describes the ability of the robot to communicate internal state to an external observer. Our hypothesis was that active learning could serve as a transparency mechanism to human teachers by highlighting areas in the space of possible examples about which the learner lacked confidence. Our preliminary results showed that such transparency was more complex than we had initially predicted. Following that pilot study, we maintained the hypothesis that a robot using active learning could achieve more accurate and faster learning as compared to a passive learner, but we also considered that using active learning could lead to problematic balance of control from an interaction perspective. In particular, we hypothesized that the naïve approach of using a constant stream of queries could be detrimental to a human teacher's enjoyment or lead to worse mental models due to lack of engagement—in some ways *decreasing* the transparency of the learner.

In this experiment, 24 human subjects each taught four concepts to our upper-torso humanoid robot Simon. The concepts corresponded to four interaction modes that we designed and implemented: *supervised learning* (SL), *active learning* (AL), *mixed initiative* (MI), and *any questions* (AQ). The SL mode is a passive supervised learner that does not use queries. We refer to the last three as *interactive modes* because they include queries in some form. AL represents a naïve active learner that queries every turn. MI is an active learner that waits for certain conditions before querying. In the AQ mode, queries are only given when solicited by the human teacher. Not surprisingly, we found that the three interactive modes outperformed supervised learning from a machine learning perspective, in both final accuracy and number of labeled examples required to achieve accurate models. Additionally, we found that people preferred the interactive modes over the passive supervised learning condition. When the robot asked questions, the subjects found it to be more intelligent, more engaging, and easier to teach. We also report several observations related to the transparency of the learning process, the balance of control in active learning, and

Manuscript received January 31, 2010; revised May 10, 2010; accepted May 12, 2010. Date of publication May 24, 2010; date of current version June 30, 2010. This work was supported by the National Science Foundation, under Grant IIS-0812106.

The authors are with the School of Interactive Computing, and the Robotics and Intelligent Machines Center, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: maya@cc.gatech.edu; cchao@cc.gatech.edu; athomaz@cc.gatech.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAMD.2010.2051030

the compliance of the human partner in answering the robot's queries. These results should inform the design of future robot active learners that are situated in human environments.

II. Background

Our approach to SG-ML is informed by child development and the human learning process. The field of situated learning studies the social world of a child and how it contributes to the child's development [2]. In a situated learning interaction, a good instructor maintains a mental model of the learner's understanding and structures the learning task with appropriate feedback and guidance. The learner contributes to the process by expressing internal state via communicative acts (e.g., asking questions or expressing understanding, confusion, and attention) [3], [4]. This is a reciprocal and tightly coupled interaction in which the learner actively influences the teacher, thereby improving his own learning environment. This situated learning process stands in contrast to typical scenarios of machine learning, which are often neither interactive nor intuitive for the human partner.

Leveraging human input has received considerable attention in both the machine learning and robotics communities. Much prior work deals with the scenario in which a machine learns by observing human behavior [5]–[7]. Other work has focused on how a machine can learn tasks from human instruction [8], [9], with human advice [10], [11], or through intervention on the learner's actions during the learning process [12].

Active learning is a relatively recent learning paradigm in which the learner chooses the examples from which it is going to learn [13]–[15]. It depends on the existence of an external oracle to provide the labels for queried examples. Active learning techniques are often evaluated in terms of the reduction in number of labels required to learn a sufficiently accurate classifier over passive supervised learning. There is overwhelming evidence that active learning can achieve significant gains in this respect, especially in applications involving text, image, and video classification where label acquisition is costly, but unlabeled data is abundant. However, most performance gain results reported in the literature are obtained with an *automated oracle* rather than an actual person [16]–[18].

Some user studies have been performed to evaluate active learning systems for estimating the cost of different detail levels of annotations or queries [19], [20], the feasibility of different types of queries (e.g., feature queries or multiple-instance queries) [21], or the effectiveness of different user interfaces while using nonexpert annotators as oracles [21], [22].

The active learning setting in the study presented here is fundamentally different from other evaluations of active learning because the active learner receives labels through a dynamic teaching–learning *interaction* rather than through monotone query labeling or annotation. In our scenario, the person is trying to teach the robot independent of whether the robot makes queries. The learner can thus, gain information without making a query, and the learner is not guaranteed that the teacher will answer its query. This scenario raises a different set of questions, such as deciding *when* to make a query in order to maximize information gain while maintaining the engagement of the teacher.

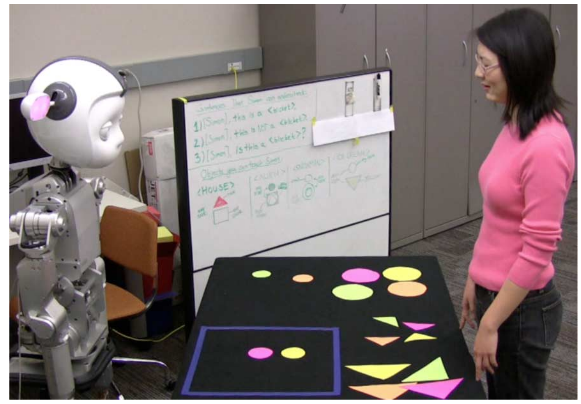


Fig. 1. Simon robot interacting with a teacher.

Research on interaction design within the human–computer interaction community is also relevant for our work. It is possible to draw parallels between some of the principles we mention in this paper and the principles proposed by Horvitz for designing mixed-initiative user interfaces [23].

There is some prior work on using active learning with humans. One system showed confidence-based active learning with (nonsocial) human labelers [24]. Similarly, Lopes *et al.* used active learning to select the states in which an expert human should be queried for an appropriate action [25]. Rosenthal *et al.* investigated how the accuracy of a human teacher's answers to a robot's questions could be improved by augmenting the questions with information about the robot's state [26].

Prior work has tended to look at the problem of human-in-the-loop active learning from the perspective of the machine learner. In this work, we also take the perspective of the human partner. This approach uncovers several issues surrounding the design of active learning systems in the domain of HRI.

III. Approach

A. Robot Platform

The robotic platform for this research is "Simon," an upper-torso humanoid social robot with two 7-DOF arms, two 4-DOF hands, and a socially expressive head and neck, including two 2-DOF ears with full RGB spectrum LEDs (Fig. 1). We are developing the Simon platform specifically for face-to–face human–robot interaction. In our task scenarios, the robot works alongside or across from a human partner at a tabletop workspace. The robot has the ability to perform simple gestures (e.g., pointing, head nods and shakes) to communicate about objects that the human can use for teaching.

Our learning system is implemented within the C6 software system (see [27]), which has a specific pipeline for triggering robot actuations from sensory inputs. In C6, a robot is equipped with various sensors such as cameras for vision and microphones for speech recognition. At every time step, the robot receives observations $O = \{o_1, \ldots, o_k\}$ from these sensory processes. The perception system is a set of percepts $P = \{p_1, \ldots, p_n\}$. Each $p \in P$ is an atomic classification and data-extraction unit that models an aspect of each observation from the sensory system by returning a match probability such that $p(o) = m$, where $m \in [0, 1]$ is a match value. The percept provides a useful level of abstraction for reducing the dimensionality of incoming sensory information.

The belief system maintains the belief set B by integrating these percepts into discrete object representations (based on spatial relationships, and various other similarity metrics or tracking mechanisms). Belief objects that detail the perceived state of the world are sent to the action system for decision making. The action system is structured as action groups of hierarchical action tuples requiring preconditions, executables, and postconditions. After a high-level action is selected, the lower level joint trajectories are transmitted to the motor module controlling the physical robot.

B. Domain Description

In this work, Simon's learning task involves colorful paper cutouts, which we refer to as *objects*. Each object has a *color* attribute with four possible values (pink, green, yellow, or orange), a *shape* attribute with three possible values (square, triangle, or circle), and a *size* attribute with two possible values (small or large), for a total of 24 possible unique objects.

Simon learns from labeled examples of object configurations. A configuration consists of a top and a bottom object and is referred to as a *compound object*. Each compound object has a total of six features: $\text{color}_{\text{top}}$, $\text{shape}_{\text{top}}$, size_{top}, $\text{color}_{\text{bottom}}$, $\text{shape}_{\text{bottom}}$, and $\text{size}_{\text{bottom}}$. Top and bottom are from Simon's perspective. The distance between objects and the orientation of objects are ignored. Simon's workspace contains exactly one of each object, so there are 552 (24×23) possible compound objects. The set of all possible compound objects is referred to as the *instance space*.

Simon's workspace is a table covered by a black tablecloth. The 24 objects are arranged in six separate groups on the perimeter of this table. Each group contains objects of the same size and shape, but of varying color. The center of the table immediately in front of Simon is the demonstration area and is demarcated to the human with a rectangular tape boundary (Fig. 1).

C. Perception

Simon expects to find exactly two objects in the demonstration area whenever a label is provided by the teacher. Objects lying on a table in front of Simon are detected through a fixed overhead camera and segmented using background subtraction. The shape of the object is recognized by the number of corners of the simplified polygon contour of the segmented object. The polygon simplification algorithm considers polygons with up to eight corners. Therefore the circle is always perceived to have eight corners, while the triangle has three and the square has four. Size is recognized based on the area within the contour and color is recognized using the color histogram of the segmented object. The objects are localized in robot world coordinates using a fixed homography from the image plane to the table plane.

In order to reduce errors, the output of the vision system is monitored by an experimenter during the interaction. If the detected configuration is inconsistent with the real configuration, the experimenter asks the subject to adjust the positions of the paper cutouts until the vision output is correct. This is usually done by moving the pieces closer to the center of the demonstration area or moving overlapping pieces away from each other. Only correct configurations are processed by the robot and logged.

The perception of speech commands is also directly controlled by the experimenter rather than using a speech recognition system in order to avoid errors. There are four possible sentences that the subject can say to the robot. The experimenter thus presses one of the four corresponding buttons on a graphical interface when the subject utters a valid sentence. The commands and the sentences are described in Section V-A. This is the only input provided by the experimenter in this experiment; everything else is autonomous.

D. Actions

Simon uses speech synthesis and gaze directions to interact with the human teacher. There are four types of actions that Simon can perform.

a) Turn Passing: Simon looks up and blinks his ears to indicate that he is ready for another example.

b) Acknowledgement: To acknowledge that the example given by the teacher has been processed, Simon gives a verbal confirmation such as "okay" or "thank you."

c) Answering: Simon uses a gesture (head nod or shake) in conjunction with speech ("Yes, this is a house." or "No, it's not.") to respond to test questions.

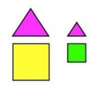
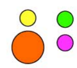
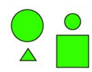
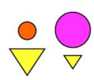
d) Making a Query: Simon can request labels for specific compound objects. He does this by requesting to replace the top or bottom piece of the compound object that is currently on the demonstration area. The request is communicated through speech synthesis using a sentence such as, "Can you replace the bottom piece with a *large pink circle*?". During this utterance, Simon also gazes towards the group on the workspace that contains the requested object (*large circles*) and changes the color of his ears to the color of the requested object (*pink*).

E. Interaction

In this study, subjects teach concepts to Simon through a turn-taking interaction. First, the teacher prepares a compound object in the demonstration area. Then, the teacher can do one of two things: 1) label the compound object's membership in the concept that is being taught as positive or negative; or 2) test Simon on the compound object by asking him to predict a label. The teacher accomplishes 1) or 2) by saying sentences from a predefined grammar. When a sentence is heard, Simon gazes towards the demonstration area and perceives the compound object. If the teacher provides a label, Simon learns from this new labeled example and then acknowledges the example. If the teacher tests Simon, Simon responds with a predicted label based on what he has learned so far. The teacher's next turn starts when Simon blinks his ears indicating that he is ready to see another example.

In the interactive modes, Simon may make a query as described in Section III-D after the person provides a labeled example. In these cases, Simon blinks his ears after finishing the query.

TABLE I
CONCEPTS

Concept Name	Concept Representation	Examples	# of Instances
HOUSE	$shape_{top} = triangle \wedge$ $color_{top} = pink \wedge$ $shape_{bottom} = square$		16
SNOWMAN	$shape_{top} = circle \wedge$ $size_{top} = small \wedge$ $shape_{bottom} = circle$		28
ALIEN	$shape_{top} = circle \wedge$ $color_{top} = green \wedge$ $color_{bottom} = green$		10
ICE CREAM	$shape_{top} = circle \wedge$ $shape_{bottom} = triangle \wedge$ $color_{bottom} = yellow$		16

IV. LEARNING

This section describes the concept learning problem and the query mechanism for active learning.

A. Concepts

The learning task in this experiment is to obtain a general compound object representation from examples of members and nonmembers of a concept. A concept is represented as a monotone conjunction of compound object attribute values that must hold true to be a member of that concept. For instance, a HOUSE is a compound object that has a pink and triangular top piece, and a square bottom piece, as shown in Table I. The size of either piece and the color of the bottom piece do not matter. Thus, the concept HOUSE is represented with the conjunction $\{\text{color}_{\text{top}} = \text{pink} \wedge \text{shape}_{\text{top}} = \text{triangle} \wedge \text{shape}_{\text{bottom}} = \text{square}\}$. This can also be denoted with $\langle \text{pink}, \text{triangle}, *, *, \text{square}, * \rangle$ where $*$ means that the value does not matter, assuming the order of features to be as given in Section III-B.

In this experiment, Simon is tasked with learning the four different concepts described in Table I. All concepts are chosen to have three terms. The actual number of possible instances for each concept differs due to the number of values that features can have and the uniqueness of objects.

B. Concept Learning

A concept learning algorithm produces a *hypothesis* based on a set of labeled examples consistent with an unknown target concept. In this experiment, the learner's hypotheses are of the same form as the target concepts described in Section IV-A, therefore exact learning is possible. The learner tries to estimate the target concept by producing a hypothesis that is consistent with the given examples. The learner chooses its hypothesis from the *hypothesis space*—the set of all possible hypotheses.

In this experiment, concepts are learned using a version space approach. The *version space* is traditionally defined as the subset of the hypothesis space that contains hypotheses consistent with all labeled examples provided to the learner [28]. In order to accommodate noisy data and errors in labeling, we quantify the *consistency* of any given hypothesis as the number of seen examples with which it is consistent, and we define the *version space* to be the set of hypotheses having the highest consistency value among all members of the hypothesis space. This relaxes the requirement that hypotheses be consistent with every previously seen labeled example, allowing the learner to recover from labeling mistakes or perception errors.

The learning algorithm updates the consistency value of each hypothesis after receiving a new example and reexpands the version space. This algorithm is summarized in Algorithm 1.[1] We provide an example in Table II to illustrate how the version space changes as labeled examples are received.

Algorithm 1 Concept learning algorithm. Where H is the set of all hypotheses, and V is the version space.

loop
  Obtain new labeled example s_i
  for $h_j \in H$ **do**
    if h_j is consistent with s_i **then**
      $consistency(h_j) \leftarrow consistency(h_j) + 1$
    end if
  end for
  Initialize $V \leftarrow \emptyset$
  $BestConsistency = \max(consistency(H))$
  for $h_j \in H$ **do**
    if $consistency(h_j) = BestConsistency$ **then**
      $V \leftarrow V \cup h_j$
    end if
  end for
end loop

C. Label Prediction

At any time during learning, the current version space can be used to predict the label of a new instance by serving as a committee. The decision is based on the majority in the predictions of the hypotheses in the version space. The confidence in the prediction is defined as the distance of the majority label to the average label. Thus, for the degenerate case of having no majority label, the confidence is 0. Table III shows examples of labels predicted by the version space committee after the learner has seen the three labeled examples shown in Table II. In the first example, all of the hypotheses in the version space agree on the label of the example, so the confidence is 1. In the second

[1]For efficiency, our implementation does not expand the complete hypothesis space, but rather considers all hypotheses that are consistent with positive examples. A generalized version of the algorithm is presented here for clarity.

TABLE II
CONCEPT LEARNING FOR HOUSE: PROGRESS OF THE VERSION SPACE AS NEW LABELED EXAMPLES ARE PROVIDED

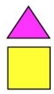
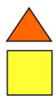
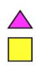
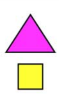
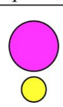
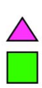
Step	Example	Label	Version Space
1	 <pink,triangle,large, yellow,square,large>	+	**Most specific hypothesis:** <pink,triangle,large yellow,square,large> **Most general hypotheses:** <pink,∗,∗,∗,∗,∗> <∗,triangle,∗,∗,∗,∗> . . . <∗,∗,∗,∗,∗,large>
2	 <orange,triangle,large, yellow,square,large>	-	**Most specific hypothesis:** <pink,triangle,large yellow,square,large> **Most general hypothesis:** <pink,∗,∗,∗,∗,∗>
3	 <pink,triangle,small, yellow,square,small>	+	**Most specific hypothesis:** <pink,triangle,∗ yellow,square,∗> **Most general hypothesis:** <pink,∗,∗,∗,∗,∗>

TABLE III
LABEL PREDICTION FOR HOUSE

Example	Version Space Predictions	Total	Prediction
 <pink,triangle,large yellow,square,small>	<pink,triangle,∗ yellow,square,∗>→ + . . . <pink,∗,∗,∗,∗,∗>→ +	+: 8 -: 0	+
 <pink,circle,large yellow,circle,small>	<pink,triangle,∗ yellow,square,∗>→ - . . . <pink,∗,∗,∗,∗,∗>→ +	+: 2 -: 6	-
 <pink,triangle,small green,square,small>	<pink,triangle,∗ yellow,square,∗>→ - . . . <pink,∗,∗,∗,∗,∗>→ -	+: 4 -: 4	?

example, six out of eight hypotheses label the example as a nonmember, so the predicted label is negative, and the confidence is smaller than 1. In the last example, there are equal votes for positive and negative, so the confidence is 0, and the predictor outputs no label.

D. Active Learning of Concepts

Active learners have a mechanism for selecting examples to be labeled by an oracle [15]. Ideally, these examples are maximally informative to the learner and thus, significantly reduce the number of required labels to create a good model.

For Simon's active learning mechanism, we implemented query-by-committee for example selection [29]. This method uses a committee of competing hypotheses and selects the example that results in maximal disagreement between the hypotheses in terms of predicted labels. The committee in our implementation is the version space. The effect of iteratively labeling examples selected by the committee is to prune away as much of the committee as possible until only one correct hypothesis is left, reducing the entropy of the committee. Note that there often exists more than a single example that results in maximal disagreement of the committee.

As an example, consider the state of the version space after seeing the three labeled examples given in Table II. As shown in Table III, some of the examples in the instance space are classified with full confidence. For other examples, the version space is split equally in terms of the predicted label—that is, half of the hypotheses are inconsistent with the example. When the version space is split in such a manner, the result of labeling the example is to halve the size of the version space. In other words, there is maximal disagreement on this example, and therefore it is a good query candidate. It is also a valid query because it differs from the last presented example by only the bottom object.

Sometimes examples cannot be queried because Simon can only request that the human change the current example by a single object, either the top or the bottom. If there are no useful queries that differ by a single object, Simon will attempt to query the preferred example in two requests. The intermediate compound object that is requested first can therefore, be uninformative for learning. We designed Simon's queries in this manner to simplify the interaction and to keep the cognitive load reasonable for the human teacher. A secondary effect of this choice is that the human teacher has the opportunity to perform better than the active learner by selecting next examples that vary both the top and the bottom objects.

V. EXPERIMENT

For this experiment, we had 24 subjects teach four concepts to the Simon robot. Each subject taught Simon in four different interaction modes. We describe the conditions and experiment protocol in this section.

A. Teaching Task

Subjects were tasked with teaching Simon the four different concepts shown in Table I in four separate teaching sessions. They were told that Simon's memory was wiped before every session and that they had to teach from scratch.

In order to teach Simon, subjects were told to arrange a compound object in the demonstration area from Simon's perspective and say one of three possible sentence types.

- [Simon], this is a ⟨concept−name⟩.
- [Simon], this is *not* a ⟨concept−name⟩.
- [Simon], is this a ⟨concept−name⟩?

Taken in conjunction with an example compound object, the first of these sentences represents a positive label, the second represents a negative label, and the third represents a test question. An experimenter used a graphical interface to submit the appropriate speech percept to Simon's perception system. Simon would then process this statement and respond to it. Subjects were instructed to listen to Simon's verbal response and to wait for Simon to blink the lights on his ears before continuing. Subjects were also told to put any pieces used for

TABLE IV
INTERACTION MODES

Mode	Concept
Supervised Learning (SL)	HOUSE
Active Learning (AL)	SNOWMAN
Mixed Initiative (MI)	ALIEN
Any Questions (AQ)	ICE CREAM

demonstrations back in their original locations in the workspace when not currently in use.

A whiteboard near Simon provided reminders about the concept representation and the types of sentences that the teacher could say. This was preferred over giving the human a piece of paper with instructions so that people would look up towards the robot rather than fixate on the piece of paper.

Subjects were instructed to continue teaching each concept until they were satisfied that Simon had learned the concept well or thought that Simon had stopped making progress.

B. Interaction Modes

We implemented four different interaction modes, SL, AL, MI, and AQ.

1) SL—The robot makes no queries during this mode. The human teacher uses positive labels, negative labels, and test questions to teach the robot. This is the baseline from which the next three modes are extended.
2) AL— Simon makes a query after every positive or negative label from the human teacher to influence the teacher's next example.
3) MI—Simon decides to make a query only under two conditions: 1) immediately following an uninformative label from the teacher; or 2) when the percentage of the instance space that is uninformative exceeds a threshold, which we set to 0.8.
4) AQ—Only during this mode, the teacher is allowed to use an additional sentence:
   - [Simon], do you have any questions?

   This sentence is meant to refer to the current example in the demonstration area. Only after this sentence is used by the teacher does Simon look at the current example and make a query.

All subjects taught Simon once in every interaction mode. Table IV shows which concept was taught for each mode.

We call the last three modes the *interactive modes*. Every subject began by teaching in the SL mode as a baseline, but the order of the interactive modes was varied. Each of the six possible orderings of the interactive modes was represented four times in our data.

Between the SL session and the first interactive session, the experimenter explained to the subject that Simon was able to request specific examples by asking for either the top or the bottom piece in the demonstration area to be switched with a different piece. The experimenter also emphasized that the subject was not required to do what Simon asked for if they did not want to. We did not ask the human teacher to comply strictly with the robot's queries because we were also interested in how well people would naturally comply with the queries in different interaction modes.

C. Survey Questions

Subjects were asked to answer questions in a web-based survey. After each teaching session, while the interaction was still fresh in the subjects' head, they were asked to answer the following questions about that specific teaching session.

- Who had more control over Simon's learning process (1–7 rating scale)?
  — 1 = You had complete control;
  — 4 = Equal control;
  — 7 = Simon had complete control.
- How well do you think Simon learned this object? Please indicate the percentage of future objects you think Simon will be able to identify correctly (0–100).
- What was your general teaching strategy for this object?
- How did you decide when this learning session should end?

After teaching in all of the interaction modes, subjects answered the following questions asking them to compare the four interaction modes.

- How difficult was it to teach each object (1–7 rating scale)?
  — 1 = Extremely easy;
  — 7 = Extremely difficult.
- While teaching each object, how clear was your mental model of what Simon knew already and what he still had to learn (1–7 rating scale)?
  — 1 = Extremely vague;
  — 7 = Extremely clear.
- How intelligent did the robot seem during each object learning session (1–7 rating scale)?
  — 1 = Extremely unintelligent;
  — 7 = Extremely intelligent.
- How enjoyable was each method of teaching (1–7 rating scale)?
  — 1 = Extremely disagreeable;
  — 7 = Extremely agreeable.
- Of the following two sessions, which teaching method did you prefer overall? (Forced choice)
  — ALIEN
  — ICE CREAM

Each of these questions included an optional comment box that was labeled, "Why? Please explain," allowing subjects to describe their thoughts in more detail. The web form also allowed subjects to page back and forth to examine or modify their answers before submission.

D. Data Logged

We consider an important event either a teacher's sentence or the robot's response. The following are the data about each event that we logged for every concept teaching session in order to characterize the differences between the interaction modes.

- *Mode*—which interaction mode was used for this session.
- *System Time*—wall clock time.
- *Interaction Step*—number of events so far for this concept.
- *Current Example*—the compound object currently configured in the demonstration area.
- *Current Label*—the label provided for the compound object, if any.
- *Sentence Type*—one of the four valid sentences types from the human teacher, if any.

TABLE V
LEARNING PERFORMANCE METRICS FROM THE ROBOT'S PERSPECTIVE

Mode	Final F_1-scores	Subjects who achieved $\|V\| = 1$	Number of Steps to reach $\|V\| = 1$
SL	77.81%	6/24 (25%)	M = 10.67, SD = 4.59
AL	100.00%	24/24 (100%)	M = 8.29, SD = 2.29
MI	98.61%	23/24 (96%)	M = 8.35, SD = 1.56
AQ	97.50%	20/24 (83%)	M = 9.30, SD = 3.53

- *Answer Type*—one of the five responses the robot could provide, if any.
- *Query*— the compound object selected for attention direction using active learning, if any.

VI. RESULTS

In this section, we describe the performance gains in the interactive modes compared to supervised learning. We also use the survey data to characterize the experience of interacting with the robot in each of the different modes.

A. Learning Performance

There already exists ample theoretical evidence supporting the use of active learning over passive supervised learning. Active learning techniques can produce classifiers with better performance using a limited number of labeled examples, or reduce the number of examples required to reach certain performance [15]. The question we have posed in this experiment is whether or not a nonexpert serving as the robot's oracle in an interactive setting can achieve such performance gains.

From the machine learning perspective, we found that using active learning significantly improved performance compared to passive SL, while we did not find a significant difference between the three different interactive modes (AL, MI, AQ) in which active learning was used. These results are summarized in Table V. This holds true across the following metrics of performance:

1) overall resulting accuracy at the end of the learning session, using the F_1-score

$$F_1 = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}$$

2) Percentage of subjects who taught the concept correctly and precisely (i.e., subjects who reached a version space consisting of a single hypothesis that was the true hypothesis for the concept).
3) Speed of convergence to the correct concept hypothesis (i.e., number of labeled examples given until the version space consisted only of the true hypothesis).

Our results confirm that active learning has better performance than passive learning, both in terms of the accuracy of the learned classifier and the number of examples required to converge to a high accuracy. In addition, our experiment showed that nonexpert subjects could successfully act as oracles when responding to a robot's queries.

Active learning proved especially useful in terms of coverage of the instance space as indicated by the low percentage of subjects who could teach the concept precisely in the SL mode (six out of 24 subjects). This shows that it can be difficult for human teachers to keep track of how much of the space they have covered even for a relatively small instance space, and that queries can significantly improve coverage.

We did not find a significant difference in performance between the three interactive modes that used active learning. Since we did not enforce compliance with the robot's queries, the achieved performance is a result of both examples elicited through queries and examples given by the subjects independent of the queries. Compliance of the subjects with the robot's queries is discussed in more detail in Section VI-E. It is possible that we would see some differentiation in performance between the interaction methods if the task were more complex and required more examples. This is a point we would like to investigate in future work. For now, we turn our attention to the question of which mode was most preferable from an interaction perspective.

B. Interactive Preferred Over Supervised

From the human partner's perspective, all of the interactive modes were more enjoyable and preferred over the SL mode. Ratings from the survey are shown in Table VI, and a selection of subjective comments are shown in Table X. Table VII gives the significance of the effect of the different interaction modes on these ratings as measured by one-way ANOVA tests, and comparison of the SL mode with the three interactive modes using t-tests.

1) Perceived Intelligence: Simon's intelligence was rated by subjects as being higher in the three interactive modes than in the SL mode. Subjects commented that Simon's questions were "relevant" and "made him seem smarter," and that he came up with "combinations that [they] had not considered or had forgotten to consider." One subject stated that Simon's ability to come up with examples himself was "definitely a sign of more intelligence than just passively listening to my instructions."

In reality, not all of the robot's queries to the teacher were informative. This was due to the nature of only requesting the top or bottom half of the example. Several subjects who perceived this shortcoming tested the robot with the queried intermediate example instead of labeling it, only to find out that the robot knew the answer already. One subject also stated that he "couldn't be sure then if [Simon] was asking for the right examples," showing his distrust of the informativeness of the queries. However, overall the subjects still took the robot's contributions to the learning process as demonstrating higher intelligence than never asking any questions at all.

2) Ease of Teaching: The fully active AL mode was considered the easiest out of the four modes. Subjects claimed that Simon's asking of questions made it "easy to know what he knows and doesn't know." One subject considered it "easier to have Simon let me know if he needed to see an example," as compared to doing all of the teaching himself in the SL mode. Another subject commented that the AL mode was "really easy because I didn't have to think."

One thing we observed in the comments was that many people rated the ease of the mode based on how conceptually intuitive the concept for that mode was to them. Although we attempted to design the concepts to be similar in terms of specificity and the number of examples required to teach them, many subjects

TABLE VI
SUBJECTIVE RATINGS (1–7)

Mode	Control (7=robot)	Difficulty (7=hardest)	Mental Model (7=clearest)	Intelligence (7=best)	Enjoyability (7=best)
SL	M = 2.50, SD = 1.38	M = 2.92, SD = 1.28	M = 4.08, SD = 1.82	M = 4.08, SD = 1.64	M = 4.92, SD = 1.79
AL	M = 4.46, SD = 1.67	M = 2.17, SD = 1.24	M = 5.08, SD = 1.77	M = 5.54, SD = 1.53	M = 5.46, SD = 1.77
MI	M = 3.25, SD = 1.45	M = 3.00, SD = 1.35	M = 5.17, SD = 1.40	M = 5.12, SD = 1.57	M = 5.46, SD = 1.50
AQ	M = 3.25, SD = 1.39	M = 2.88, SD = 1.57	M = 5.17, SD = 1.55	M = 5.75, SD = 1.29	M = 6.04, SD = 1.16

TABLE VII
SIGNIFICANCE TESTS ACROSS ALL MODES (REPEATED MEASURES ANOVA) AND COMPARISON OF THE SL MODE WITH THE THREE INTERACTIVE MODES ON SURVEY RATINGS

Test	Control (7=robot)	Difficulty (7=hardest)	Mental Model (7=clearest)	Intelligence (7=best)	Enjoyability (7=best)
ANOVA	$F(3,23)=7.69$, $p<.001$*	$F(3,23)=2.69$, $p>.05$	$F(3,23)=3.39$, $p<.05$*	$F(3,23)=13.02$, $p<.001$*	$F(3,23)=4.57$, $p<.01$*
SL vs. AL	$t(23)=-4.15$, $p<.001$*	$t(23)=2.16$, $p<.05$*	$t(23)=-2.06$, $p>.05$	$t(23)=-5.05$, $p<.001$*	$t(23)=-2.01$, $p>.05$
SL vs. MI	$t(23)=-1.94$, $p>.05$	$t(23)=0.10$, $p>.05$	$t(23)=-2.57$, $p<.05$*	$t(23)=-5.95$, $p<.001$*	$t(23)=-3.58$, $p<.005$*
SL vs. AQ	$t(23)=-1.72$, $p>.05$	$t(23)=-0.25$, $p>.05$	$t(23)=-2.41$, $p<.05$*	$t(23)=-2.99$, $p<.01$*	$t(23)=-1.88$, $p>.05$

TABLE VIII
PERCEPTION OF ACCURACY

Mode	F_1-score	Estimated	Discrepancy
SL	77.81%	88.08%	–10.27%
AL	100.00%	95.33%	+4.67%
MI	98.61%	94.33%	+4.28%
AQ	97.50%	91.38%	+6.12%

TABLE IX
EFFICIENCY

Mode	Time	Steps	Efficiency	Informative
SL	6.82 mins	19.79 steps	5.12%/step	75.38%
AL	5.48 mins	20.42 steps	5.48%/step	79.51%
MI	5.70 mins	20.96 steps	5.17%/step	78.60%
AQ	6.30 mins	20.25 steps	5.42%/step	77.36%

perceived the complexity of the concepts as varying for the different concepts. The perceived relative complexity also varied across subjects. We thought that this could be a reason that the ratings did not show a significant difference between the modes in the way that the comments did.

3) Enjoyability: Despite a number of negative comments about each of the interactive modes (see Table X), all three interactive modes were ranked higher than the SL mode in enjoyability. AQ was the favorite among many subjects. A t-test between AQ and SL shows a significant preference for AQ, as shown on Table VII. Subjects said it was "fun to answer his questions," and that it was "extremely enjoyable getting feedback from Simon." This was in contrast to the SL mode, which one subject even described as "very dull."

C. Transparency and Active Learning

One of our hypotheses going into the experiment was that active learning could potentially help the human partner maintain a better mental model of the learning process. To measure this, we had subjects report their perceived accuracy after each learning session and examined the discrepancy between their estimates and the actual accuracy of their taught models. Results are shown in Table VIII.

People had a more accurate performance estimate in the interactive modes than in the SL mode. In addition to having a larger error in estimating the performance in the SL mode, subjects overestimated rather than underestimated the accuracy of the learner. Overestimating the performance is dangerous from a machine learning perspective because it leads to early stopping, preventing the learner from seeing more examples.

The underestimate for the interactive modes is still not ideal. In the optimal case, the teacher has a mental model that matches the accuracy of the learner and stops teaching exactly when the accuracy hits 100%. In the interactive modes, the subjects continued teaching even after the learner was done and still did not feel they were finished. As a result, there was no significant reduction in time taken, and there were no significant gains in efficiency, as shown in Table IX. In addition, the number of uninformative examples given was actually roughly equal between all four modes.

If we suppose that the SL sessions continued until the robot learned a model of 100% accuracy, then the SL sessions should have taken more time than the interactive sessions. We think that it is possible to realize gains in efficiency by reducing the time spent teaching after the robot is done learning. However, this will require the learner to demonstrate increased transparency about what has been learned or when to stop teaching.

D. Balance of Control

As shown in Table VI, the rankings people gave for balance of control are as we hypothesized. From the order of most to least human control are SL, AQ, MI, and AL. The data also show a significant effect of the constant stream of queries in the AL mode on sense of control, as shown in Table XI. Subjects rated the AL mode to have significantly less human control than all other modes and the significance level decreased in the order of SL, AQ, and MI.

Overall, we observed that the subjects' responses were skewed towards human control (closer to 1 than 7). That is, even in the AL case when the robot attempted to direct all of the examples, people reported approximately equal balance of control with an average rating of 4.46. However, several participants mentioned feeling that they held a peripheral role in the learning process during the AL condition, as shown in the negative comments in Table X. Letting the robot control the interaction can cause a teacher to stop maintaining a mental model of the process, preventing the learner from achieving maximum efficiency when learning from a good teacher.

The two hybrid approaches of MI and AQ were an attempt to bring more balance of control to the interaction without

TABLE X
SELECTED COMMENTS ON INTERACTIVE MODES

Mode	Positive comments	Negative comments
AL	"I didn't have to think" "greatly simplified the learning" "encouraging to observe his rate of learning"	"I didn't feel like there was real interaction" "asked questions several times when I didn't want to be asked" "made me feel unnecessary" "lost track of what I was trying to teach" "turned my brain off" "constantly asked questions, it was annoying"
MI	"more interactive" "more of a dialog" "required less of my time"	"wasn't confident when I was done teaching" "frustrating" "hard for me to keep straight" "confusing to determine whether he had learnt enough" "distracted by the robot"
AQ	"easier to work with" "richer interaction" "felt very comfortable... felt complete" "enjoyed being able to direct the process more" "very clear when Simon has no questions" "easier than being asked questions more often"	"less easy because I had to prompt Simon" "took more effort on my part than his" "required substantial teaching on my side"

TABLE XI
T-TESTS FOR AL MODE FOR RATINGS OF BALANCE OF CONTROL

Modes	Result
AL vs. SL	$t(23)=4.15$, $p < .001$*
AL vs. AQ	$t(23)=3.03$, $p < .01$*
AL vs. MI	$t(23)=3.14$, $p < .005$*

TABLE XII
COMPLIANCE ON ANSWERING QUERIES

Mode	Average number of queries	Percent answered
AL	6.46	78.82%
MI	3.04	68.61%
AQ	3.29	65.55%

sacrificing the benefits offered by active learning. MI represents robot-directed shared control, and AQ represents human-directed shared control. When we presented a forced choice between MI and AQ, a majority of subjects chose the AQ mode, but the results definitely show two categories of people. 33.33% of the subjects preferred the MI mode because it required less effort and seemed more balanced. The rest preferred the increased control over teaching offered by requiring permission to ask questions.

The comments shown in Table X are what we feel is a representative selection of those given in the open-ended response boxes. Positive comments about AL were characterized by the triviality of teaching and the speed of Simon's learning, and negative comments described displeasure at being bombarded by questions or having a less significant role. Positive comments about MI cited a balanced and efficient interaction, while negative comments cited confusion at the seeming randomness of the robot queries. The AQ mode had the least negative comments, which mentioned that explicitly allowing Simon to ask questions took longer, and forgetting to allow him to ask questions made the mode too much like the SL mode. The many positive AQ comments described the interaction as feeling the most natural and easy, and enjoying the control of teaching in conjunction with the support of Simon's queries.

The results seem to show that people generally prefer that the robot take initiative and be curious about the topic, but that people also tend to desire control. Thus, they prefer to direct the shared control, at least when they are doing the teaching.

When deploying active learning systems in the context of HRI, the optimal interaction strategy may be user dependent. The AQ mode seems appropriate when the teacher is an expert, for it gives the teacher ample control while still allowing him to draw assistance from the robot when necessary. Given our relatively simple domain, this may be why AQ was the most preferred. The MI mode is more appropriate for nonexperts who would tend to give uninformative examples, so it is potentially problematic if they dislike teaching using this mode of interaction. For everyday people who are teaching, either robot-directed shared control needs to be improved, or AL should be used to reduce confusion and ambiguity. Such a system may need multiple strategies for interaction, with an arbitration scheme to determine which is appropriate for the current teacher. This is an important direction for future work in this domain.

E. Compliance

Compliance describes how likely the teachers are to answer the active learner's queries. This is an issue that to our knowledge has not been addressed in the literature, but will be highly important for active learning systems that need to learn from everyday people. In order to reap the benefits of active learning, the robot needs the human teacher to answer the questions it asks.

As a measure of compliance in answering Simon's queries, we examine the percentage of queries made by Simon that were labeled by the human teacher in the next turn. We expected to see the relaxed versions of active learning (MI and AQ) lead to better compliance due to being less demanding and leaving more control to the teacher compared to traditional active learning (AL). However, the results refuted this hypothesis; we found that teachers responded to a higher fraction of the AL queries compared to the MI or AQ queries (Table XII).

We believe that this finding could be due to two issues. As mentioned previously, some of the queries made by the robot were uninformative and could already be classified by the robot. These queries were nevertheless necessary as intermediate steps to attain an informative query. The percentage of queries that are uninformative tends to be higher in the relaxed versions of

TABLE XIII
COMPLIANCE ON ANSWERING QUERIES ACCORDING TO INFORMATIVENESS OF THE QUERY

Mode	Informative queries		Uninformative queries	
	Answered	Not answered	Answered	Not answered
AL	78.23%	14.76%	0.60%	6.42%
MI	64.10%	9.51%	4.51%	21.88%

active learning. In the MI mode, the robot waits to make a query until the number of informative queries in the instance space becomes more sparse. When very few examples are informative for pruning the version space, the probability that the example on the demonstration area has one common object with one of the informative examples is low.

Table XIII compares compliance on informative and uninformative queries. These results are consistent with our original expectation: a higher percentage of *informative* queries are declined by the teacher in the AL mode. Most of the queries that were declined in the MI mode were *uninformative*. These results also show that subjects were successful in detecting uninformative queries, either by maintaining a good mental model of what the robot knew, or by testing the robot with the queried examples. Note that the occurrence of uninformative queries could also have made the teacher more reluctant to respond to informative queries.

The other potential issue has to do with the human's teaching strategy. In both MI and AQ, the lack of queries initiated by the robot early in the learning process seemed to establish a balance of control dominated by the teacher. A human teacher who seized control early in the interaction may be more reluctant to give it up later on. The teacher might be following a certain strategy, and even when informative, a robot's query could seem out of place. This points to an important area for future research: in what way should the robot ask smart questions? It may be that the "best" query is not just an informative query from the machine learning perspective, but also one that is appropriate with respect to the teacher's strategy. The robot may be able to model the teaching strategy in order to form the most *appropriate* informative queries.

Based on these observations, our design recommendations to increase compliance are 1) to avoid making uninformative queries that could weaken the teacher's trust in the utility of answering the robot's queries, and 2) to establish or relinquish control early on and subsequently maintain these mutual expectations.

VII. CONCLUSION

Our research is in SG-ML: designing algorithms and systems that take advantage of the ways that everyday people approach the task of teaching in order to build robots that learn new tasks and skills from end users. In this work, we have focused on issues surrounding the development of active learning systems in the context of human–robot interaction.

In our experiment, 24 people taught simple object-based concepts to a social robot. We evaluated three interaction modes using active learning compared with a baseline interaction using only supervised learning. Our results showed that human-controlled active learning could achieve performance gains over pure supervised learning, and that the use of active learning is viewed as preferable along with a number of dimensions by human teachers, including perception of robot intelligence, ease of teaching, and enjoyability.

Our results show that people had a more accurate performance estimate in the interactive modes using active learning than in the passive supervised learning mode, and that they tended to underestimate rather than overestimate the learner's performance. However, an issue for future work is developing transparency mechanisms to better communicate to the human teacher about when learning has been completed.

In the context of HRI, compliance will be an important issue for active learning systems. Our observations indicate that two key elements for achieving higher compliance from human teachers may be establishing roles and balance of control early on in the learning process, and avoiding uninformative queries.

In general, our experiment found a range of preferences across our three active learning modes. This makes us believe that the optimal strategy is likely to be user-dependent. The AL condition was found by some to be trivially easy and efficient, but some users felt bombarded by questions and less engaged in the process. The MI condition was seen as more balanced, but the intermittent queries were also sometimes confusing to the teachers. The AQ condition was the most preferred in our experiment. The interaction was described as the natural and easy, but was also seen as potentially less efficient than the other two active learning conditions.

Our overall result shows that people prefer that the robot take initiative as a learner and are willing, and even pleased, to accept assistance from the learner, but that they also have a desire to control that assistance. For nonexpert teachers, a robot using active learning can have a significant contribution to performance, but can also be frustrating to teachers. A challenge for future work is designing robot active learners that can achieve theoretical performance gains with nonexpert human–teachers without usurping control of the interaction.

REFERENCES

[1] C. Chao, M. Cakmak, and A. L. Thomaz, "Transparent active learning for robots," in *Proc. ACM Conf. Human-Robot Interaction (HRI)*, 2010, pp. 317–324.

[2] L. S. Vygotsky, , M. Cole, Ed., *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard Univ. Press, 1978.

[3] R. Krauss, Y. Chen, and P. Chawla, "Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us?," in *Advances in Experimental Social Psychology*, M. Zanna, Ed. Orlando, FL: Academic, 1996, pp. 389–450.

[4] M. Argyle, R. Ingham, and M. McCallin, "The different functions of gaze," *Semiotica*, vol. 7, pp. 10–32, 1973.

[5] E. Horvitz, J. Breese, D. Heckerman, D. Hovel, and K. Rommelse, "The Lumiere project: Bayesian user modeling for inferring the goals and needs of software users," in *Proc. 14th Conf. Uncertainty Artif. Intell.*, Madison, WI, Jul. 1998, pp. 256–265.

[6] Y. Kuniyoshi, M. Inaba, and H. Inoue, "Learning by watching: Extracting reusable task knowledge from visual observation of human performance," *IEEE Trans. Robot. Autom.*, vol. 10, no. 6, pp. 799–822, Dec. 1994.

[7] P. Abbeel and A. Ng, "Apprenticeship learning via inverse reinforcement learning," in *Proc. 21st Int. Conf. Mach. Learn.*, Banff, AB, Canada, 2004, pp. 1–8.

[8] M. N. Nicolescu and M. J. Matarić, "Natural methods for robot task learning: Instructive demonstrations, generalization and practice," in *Proc. 2nd Int. Conf. Autonom. Agents Multi-Agent Syst.*, Melbourne, Australia, Jul. 2003.

[9] C. Breazeal, A. Brooks, J. Gray, G. Hoffman, J. Lieberman, H. Lee, A. L. Thomaz, and D. Mulanda, "Tutelage and collaboration for humanoid robots," *Int. J. Humanoid Robot.*, vol. 1, no. 2, pp. 315–348, 2004.

[10] R. Maclin, J. Shavlik, L. Torrey, T. Walker, and E. Wild, "Giving advice about preferred actions to reinforcement learners via knowledge-based kernel regression," in *Proc. 20th Nat. Conf. Artif. Intell.*, Pittsburgh, PA, Jul. 2005, pp. 819–824.

[11] A. L. Thomaz and C. Breazeal, "Teachable robots: Understanding human teaching behavior to build more effective robot learners," *Artif. Intell. J.*, vol. 172, pp. 716–737, 2008.

[12] W. D. Smart and L. P. Kaelbling, "Effective reinforcement learning for mobile robots," in *Proc. IEEE Int. Conf. Robot. Autom.*, Washington DC, 2002, pp. 3404–3410.

[13] D. Angluin, "Queries and concept learning," *Mach. Learn.*, vol. 2, pp. 319–342, 1988.

[14] D. A. Cohn, Z. Ghahramani, and M. I. Jordan, "Active learning with statistical models," *J. Artif. Intell. Res.*, vol. 4, pp. 129–145, 1996.

[15] B. Settles, Active Learning Literature Survey Univ. Wisconsin, Madison, WI, Tech. Rep. 1648, 2010.

[16] S. Tong and D. Koller, "Support vector machine active learning with applications to text classification," in *Proc. 17th Int. Conf. Mach. Learn.*, Stanford, CA, 2000, pp. 999–1006.

[17] D. A. Cohn, R. Caruana, and A. McCallum, Semi-Supervised Clustering With User Feedback Cornell Univ., Tech. Rep. TR2003-1892, 2003.

[18] A. Culotta and A. McCallum, "Reducing labeling effort for structured prediction tasks," in *Proc. 20th Nat. Conf. Artif. Intell.*, Pittsburgh, PA, 2005, pp. 746–751.

[19] B. Settles, M. Craven, and L. Friedland, "Active learning with real annotation costs," in *Proc. NIPS Workshop Cost-Sensitive Learn.*, 2008, pp. 1069–1078.

[20] S. Vijayanarasimhan and K. Grauman, "Whats it going to cost you? predicting effort vs. informativeness for multi-label image annotations," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Miami, FL, 2009, pp. 2262–2269.

[21] G. Druck, B. Settles, and A. McCallum, "Active learning by labeling features," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, Emnlp, Singapore, 2009, pp. 81–90.

[22] A. Culota, T. Kristjansson, A. McCallum, and P. Viola, "Corrective feedback and persistent learning for information extraction," *Artif. Intell.*, vol. 170, pp. 1101–1122, 2006.

[23] E. Horvitz, "Principles of mixed-initiative user interfaces," in *Proc. ACM Conf. Human Factors Comput. Syst.*, Pittsburgh, PA, 1999, pp. 159–166.

[24] S. Chernova and M. Veloso, "Confidence-based policy learning from demonstration using Gaussian mixture models," in *Proc. Int. Conf. Autonom. Agents Multi-Agent Syst.*, Honolulu, HI, 2007.

[25] M. Lopes, F. Melo, and L. Montesano, "Active learning for reward estimation in inverse reinforcement learning," in *Machine Learning and Knowledge Discovery in Databases*. Berlin, Germany: Springer-Verlag, 2009, pp. 31–46.

[26] S. Rosenthal, A. Dey, and M. Veloso, "How robots' questions affect the accuracy of the human responses," in *Proc. IEEE Symp. Robot Human Interactive Commun.*, Toyama, Japan, 2009, pp. 1137–1142.

[27] B. Blumberg, M. Downie, Y. Ivanov, M. Berlin, M. P. Johnson, and B. Tomlinson, "Integrated learning for interactive synthetic characters," in *Proc. ACM Int. Conf. Comput. Graphics Interactive Tech.*, San Antonio, TX, 2002, vol. 21, no. 3, pp. 417–426.

[28] N. Littlestone, "Learning quickly when irrelevant attributes abound: A new linear-threshold algorithm," *Mach. Learn.*, vol. 2, pp. 285–318, 1988.

[29] H. Seung, M. Opper, and H. Sompolinsky, "Query by committee," in *Proc. ACM Workshop Comput. Learn. Theory*, 1992, pp. 287–294.

Maya Cakmak received the B.Sc. degree in electrical and electronics engineering, and the M.Sc. degree in computer engineering, both from Middle East Technical University, Turkey. She is currently pursuing the Ph.D. degree in robotics at the Georgia Institute of Technology, Atlanta.

Her research interests include human–robot interaction, social learning, and interactive machine learning.

Crystal Chao received the B.Sc. degree in computer science from the Massachusetts Institute of Technology, Cambridge, in 2008. She is currently pursuing the Ph.D. degree in robotics at the Georgia Institute of Technology, Atlanta.

Her research focuses on human–robot interaction and machine learning.

Andrea L. Thomaz received the B.Sc. degree from the University of Texas, Austin, in 1999, and the M.Sc. and Ph.D. degrees from the Massachusetts Institute of Technology, Cambridge, in 2002 and 2006, respectively.

She is currently an Assistant Professor at the School of Interactive Computing, Georgia Institute of Technology, Atlanta, where she directs the Socially Intelligent Machines Lab. Her work focuses on social robots and socially guided machine learning. She is published in the areas of robotics, human–robot interaction, and artificial intelligence.