Available online at www.sciencedirect.com

SCIENCE DIRECT®

Robotics and Autonomous Systems 54 (2006) 385–393

Robotics and Autonomous Systems

www.elsevier.com/locate/robot

Using perspective taking to learn from ambiguous demonstrations

Cynthia Breazeal*, Matt Berlin, Andrew Brooks, Jesse Gray, Andrea L. Thomaz

Media Laboratory, Massachusetts Institute of Technology, 20 Ames St., Cambridge, MA 02139, USA

Abstract

This paper addresses an important issue in learning from demonstrations that are provided by "naïve" human teachers—people who do not have expertise in the machine learning algorithms used by the robot. We therefore entertain the possibility that, whereas the average human user may provide sensible demonstrations from a human's perspective, these same demonstrations may be insufficient, incomplete, ambiguous, or otherwise "flawed" from the perspective of the training set needed by the learning algorithm to generalize properly. To address this issue, we present a system where the robot is modeled as a *socially engaged* and *socially cognitive* learner. We illustrate the merits of this approach through an example where the robot is able to correctly learn from "flawed" demonstrations by taking the visual perspective of the human instructor to clarify potential ambiguities.

Keywords: Human–robot interactions; Socially guided learning; Learning by demonstration; Perspective taking

1. Introduction

New applications for robots in the human environment motivate the need to design robots that can learn from the average consumer. To address this issue, and inspired by the way people and animals learn from others, researchers have begun to investigate various forms of social learning and interactive training techniques, e.g., imitation-based learning [15], clicker training [2], learning by demonstration [14], and tutelage [3].

This paper addresses an important issue in building robots that can successfully learn from demonstrations that are provided by "naïve" human teachers who do not have expertise in the learning algorithms used by the robot. As a result, the teacher may provide sensible demonstrations from a human's perspective; however, these same demonstrations may be insufficient, incomplete, ambiguous, or otherwise "flawed" from the perspective of providing a correct and sufficiently complete training set needed by the learning algorithm to generalize properly.

To address this issue, we present a system where the robot is modeled as a *socially engaged* and *socially cognitive* learner. First, to build a socially engaged learner, we are inspired by human tutelage where the teaching–learning process is conducted as a tightly interactive collaboration. For instance, the teacher may first demonstrate a skill as the learner observes. Then, the learner attempts to master the skill as the teacher observes. Expressive feedback on the part of the learner is important to the teacher. The teacher uses this feedback to model the mental state of the learner—e.g. confusion, frustration, curiosity, etc. Consequently, the teacher follows with a refined demonstration specifically tailored to give meaningful feedback to the learner based on the learner's demonstration and expressive feedback. In this way, the teacher guides the learner's exploration *and* the learner guides the teacher's instruction to make it more relevant and timely. We believe modeling this kind of socially engaged process on robots will support a natural, efficient, and understandable teaching experience for the human. Further, we believe that this style of interaction will contribute to keeping the human teacher engaged and motivated to teach the robot.

Second, we believe that robots will need to be socially cognitive learners that can infer the intention of the human's instruction, even if the teacher's demonstrations are less than perfect for the robot. In this case, the robot should clarify the problematic demonstration if necessary. In the spirit of the common saying "Do as I say, not as I do", this capability would allow a robot to "Learn what I mean to teach".

* Corresponding author.
E-mail address: cynthiab@media.mit.edu (C. Breazeal).

0921-8890/$ - see front matter
doi:10.1016/j.robot.2006.02.004

Meltzoff [13] showed that human infants begin to exhibit the ability to learn from flawed demonstrations at 18 months of age when they can imitate intended acts. For instance, the experimenter presented 18 month old infants with a novel object, a dumbbell, and attempted to show the infants how to remove one of the ends. Each demonstration by the experimenter is flawed in that his hand slips off the end as he tries to pull it apart. However, when the dumbbell is given to the infant, she correctly pulls the end of the dumbbell off, thereby interpreting the experimenter's demonstrations in intentional goal-directed terms. In addition, the infant will even explore novel ways to remove the end if simply pulling it off is difficult (for instance, if the dumbbell is too large for her to hold).

The outline of this paper is as follows. First, we describe our experimental platform, an expressive humanoid robot called Leonardo, and a virtual simulator for the robot. Next, we present our implementation of the robot's socio-cognitive skills and the biological inspiration that guides our approach. These skills allow the robot to not only recognize the actions of the human demonstrator, but also to infer the human's goal and belief states. As in the example of human infants imitating intended actions, Leonardo interprets human demonstrations in intentional terms. We then describe our tutelage-inspired approach to task and goal learning which positions the robot as a socially engaged learner. We present the implementation of the robot's learning mechanisms and communication skills. Finally, we present how these two capabilities, perspective taking and task learning, are integrated to model the robot as a socially engaged and socially cognitive learner, and illustrate the robot's ability to learn from "flawed" demonstrations.

2. Platform

The implementation presented in this paper runs on the Leonardo robot (Leo), a 65 degree of freedom humanoid robot (Fig. 1). Leo sees the world through two environmentally mounted stereo-vision cameras. One stereo camera is mounted behind Leo's head for detecting humans within the robot's interpersonal space. The second stereo camera looks down from above, and detects objects in Leo's space as well as human hands pointing to these objects. Leo can use his eye cameras for fine corrections to look directly at objects or faces at a higher resolution. In order to perceive the upper torso pose of a human interacting with the robot, Leo uses the VTracker articulated body tracking software developed and generously provided for our use by the Vision Interfaces group at the MIT Computer Science and Artificial Intelligence Laboratory.

3. Socio-cognitive skills

Our approach to endowing machines with socially-cognitive learning abilities is inspired by leading psychological theories and recent neuroscientific evidence for how human brains might infer the mental states of others and the role of imitation as a critical precursor. Specifically, *Simulation Theory* holds that certain parts of the brain have dual use; they are used to not only generate our own behavior and mental states, but also to predict and infer the same in others. To understand another

Fig. 1. The Leo robot and simulator.

person's mental process, we use our own similar brain structure to simulate the introceptive states of the other person [7,8,1].

Inspired by this theory, our simulation-theoretic approach and implementation enables a humanoid robot to monitor an adjacent human teacher by simulating his or her behavior within the robot's own generative mechanisms on the motor, goal-directed action, and perceptual-belief levels. Due to space limitations, we focus our technical presentation on the design of the robot's perceptual-belief systems and the simulation-theoretic mechanisms that reside within it. This grounds the robot's information about the teacher in the robot's own systems, allowing it to make inferences about the human's likely beliefs in order to better understand the intention behind the teacher's demonstrations. We refer the interested reader to [10] for technical details for how our simulation-theoretic mechanisms are applied to the motor system to enable the robot to recognize human action, and to the goal-directed behavior system to infer human intention.

We believe that maintaining mutual beliefs and common ground in human-robot teaching–learning scenarios will make robots more efficient and understandable learners, as well as more robust to the miscommunications or misunderstandings that inevitably arise even in human–human tutelage [4].

Specifically, in our demonstration, the robot learns the intended task through a tutelage-style interaction. As the robot observes the human's demonstrations, it internally simulates "what might I be trying to achieve were I performing these demonstrations in their context?" The robot therefore interprets and hypothesizes the intended concept being taught not only from its own perspective, but from the human teacher's visual perspective as well. Through this process, the robot successfully identifies ambiguous demonstrations given by the human instructor, and clarifies the human's intent behind these confusing demonstrations. Once these problematic demonstrations are disambiguated, the robot correctly learns the intended task.

4. Belief modeling

This section presents a technical description of two important components of our cognitive architecture: the Perception System and the Belief System. The Perception System is responsible for extracting perceptual features from raw sensory information, while the Belief System is responsible for integrating this information into discrete object representations. The Belief System represents our approach to sensor fusion, object tracking and persistence, and short-term memory.

On every time step, the robot receives a set of sensory observations $O = \{o_1, o_2, \ldots, o_N\}$ from its various sensory processes. As an example, imagine that the robot receives information about buttons and their locations from an eye-mounted camera, and information about the button indicator lights from an overhead camera. On a particular time step, the robot might receive the observations $O = \{$(red button at position (10,0,0)), (green button at (0,0,0)), (blue button at (−10,0,0)), (light at (10,0,0)), (light at (−10,0,0))$\}$.

Information is extracted from these observations by the Perception System. The Perception System consists of a set of *percepts* $P = \{p_1, p_2, \ldots, p_K\}$, where each $p \in P$ is a classification function defined such that

$$p(o) = (m, c, d), \tag{1}$$

where $m, c \in [0, 1]$ are match and confidence values and d is an optional derived feature value. For each observation $o_i \in O$, the Perception System produces a *percept snapshot*

$$s_i = \{(p, m, c, d) \mid p \in P, p(o_i) = (m, c, d), m * c > k\}, \tag{2}$$

where $k \in [0, 1]$ is a threshold value, typically 0.5. Returning to our example, the robot might have four percepts relevant to the buttons and their states: a location percept that extracts the position information contained in the observations, a color percept, a button shape recognition percept, and a button light recognition percept. The Perception System would produce five percept snapshots corresponding to the five sensory observations, containing entries for relevant matching percepts.

These snapshots are then clustered into discrete object representations called *beliefs* by the Belief System. This clustering is typically based on the spatial relationships between the various observations, in conjunction with other metrics of similarity. The Belief System maintains a set of beliefs B, where each belief $b \in B$ is a set mapping percepts to history functions: $b = \{(p_x, h_x), (p_y, h_y), \ldots\}$. For each $(p, h) \in b$, h is a history function defined such that

$$h(t) = (m'_t, c'_t, d'_t) \tag{3}$$

represents the "remembered" evaluation for percept p at time t. History functions may be lossless, but they are often implemented using compression schemes such as low-pass filtering or logarithmic timescale memory structures. For convenience, we define the attribute function

$$a(b, p) = \begin{cases} 1 & \text{if } (p, h) \in b \text{ for some } h \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

which indicates whether or not a belief contains a history function for a particular percept.

A Belief System is fully described by the tuple (B, G, M, d, q, w, c, v), where

- B is the current set of beliefs,
- G is a generator function map, $G : P \to \mathcal{G}$, where each $g \in \mathcal{G}$ is a history generator function where $g(m, c, d) = h$ is a history function as above,
- M is a merge function map, $M : P \to \mathcal{M}$, where each $m \in \mathcal{M}$ is a history merge function where $m(h_1, h_2) = h'$ represents the "merge" of the two histories h_1 and h_2,
- $d = d_1, d_2, \ldots, d_L$ is a vector of belief distance functions, $d_i : B \times B \to \mathcal{R}$,
- $q = q_1, q_2, \ldots, q_L$ is a vector of indicator functions, where each element q_i denotes the applicability of d_i, $q_i : B \times B \to \{0, 1\}$,
- $w = w_1, w_2, \ldots, w_L$ is a vector of weights, $w_i \in \mathcal{R}$,
- $c = c_1, c_2, \ldots, c_J$ is a vector of culling functions, $c_j : B \times B \to \{0, 1\}$, and
- v is a scalar threshold value.

Using the above, we define the Belief Distance Function D, the Belief Merge Function R, and the Belief Culling Function C:

$$D(b_1, b_2) = \sum_{i=1}^{L} w_i q_i(b_1, b_2) d_i(b_1, b_2) \tag{5}$$

$$\begin{aligned} R(b_1, b_2) = b' = {} & \{(p, h) \mid (p, h) \in b_1, a(b_2, p) = 0\} \\ & \cup \{(p, h) \mid (p, h) \in b_2, a(b_1, p) = 0\} \\ & \cup \{(p, h') \mid (p, h_1) \in b_1, (p, h_2) \in b_2, \\ & m = M(p), h' = m(h_1, h_2)\} \end{aligned} \tag{6}$$

$$C(b) = \prod_{j=1}^{J} c_j(b). \tag{7}$$

The Belief System manages three key processes: creating new beliefs from incoming percept snapshots, merging sets of beliefs, and culling stale beliefs. For the first of these processes, we define the function N, which creates a new belief b_i from a percept snapshot s_i:

$$\begin{aligned} b_i = N(s_i) = {} & \{(p, h) \mid (p, m, c, d) \in s_i, \\ & g = G(p), h = g(m, c, d)\}. \end{aligned} \tag{8}$$

For the second process, merging sets of beliefs, the Belief System reduces sets of beliefs by clustering proximal beliefs, assumed to represent different observations of the same object. This is accomplished via bottom-up, agglomerative clustering as follows. For a set of beliefs B:

1: **while** $\exists b_x, b_y \in B$ such that $D(b_x, b_y) < v$ **do**
2: find $b_1, b_2 \in B$ such that $D(b_1, b_2)$ is minimal
3: $B \leftarrow B \cup \{R(b_1, b_2)\} \setminus \{b_1, b_2\}$
4: **end while**

We label this process $merge(B)$.

In the third and final process, the Belief System culls stale beliefs by removing all beliefs from the current set for which $C(b) = 1$. In summation, then, a complete Belief System update cycle proceeds as follows:

1: begin with current belief set B
2: receive percept snapshot set S from the Perception System
3: create incoming belief set $B_I = \{N(s_i) \mid s_i \in S\}$
4: merge: $B \leftarrow merge(B \cup B_I)$
5: cull: $B \leftarrow B \setminus \{b \mid b \in B, C(b) = 1\}$.

Returning again to our example, the Belief System might specify a number of relevant distance metrics, including a measure of Euclidean spatial distance along with a number of metrics based on symbolic feature similarity. For example, a symbolic metric might judge observations that are hand-shaped as distant from observations that are button-shaped, thus separating these observations into distinct beliefs even if they are collocated. For our example, the merge process would produce three beliefs from the original five observations: a red button in the ON state, a green button in the OFF state, and a blue button in the ON state.

5. Belief inference and visual perspective simulation

When demonstrating a task to be learned, it is important that the context within which that demonstration is performed is the same for the teacher as it is for the learner. However, in complex and dynamic environments, it is possible for the instructor's beliefs about the context surrounding the demonstration to diverge from those of the learner.

For example, a visual occlusion could block the teacher's viewpoint of a region of a shared workspace (but not that of the learner) and consequently lead to ambiguous demonstrations where the teacher does not realize that the visual information of the scene differs between them. For instance, consider a scenario where the teacher wants to demonstrate how to assemble a structure. If the teacher incorrectly indicates that the task has been completed without realizing that part of the structure remains unfastened, due to having an obstructed view, the learner may think that fastening that part of the structure is not relevant to the task. In a teaching scenario, these kinds of issues may make the human instructor's demonstrations "flawed" from the learner's perspective and lead to learning the wrong thing.

To address this issue, Leo must establish and maintain mutual beliefs with the human instructor about the shared context surrounding demonstrations. Leo keeps track of his own beliefs about object state using his Belief System, described in Section 4. In order to model the beliefs of the human instructor as separate and potentially different from his own, Leo re-uses the mechanism of his own Belief System. Beliefs that represent Leo's model of the human's beliefs are in the same format as his own, but are maintained separately so that Leo can compare differences between his beliefs and the human's beliefs.

As described in Section 4, belief maintenance consists of incorporating new sensor data into existing knowledge of the world. Leo's sensors are all in his own reference frame, so objects in the world are perceived relative to his position and orientation. In order to model the beliefs of the human, Leo re-uses the same mechanisms used for his own belief modeling, but first transforms the data into the reference frame of the human (see Fig. 2). Leo can also filter out incoming data that he believes is not perceivable to the human, preventing that new data from updating the human's beliefs. As you recall, the sensory observations $O = \{o_1, o_2, \ldots, o_N\}$ are the input to the robot's Belief System. The inputs to the secondary Belief System which models the human's beliefs are O', where:

$$O' = \{P(o') \mid o' \in O, V(o') = 1\} \tag{9}$$

where:

$$V(x) = \begin{cases} 1 & \text{if } x \text{ is visible to human} \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

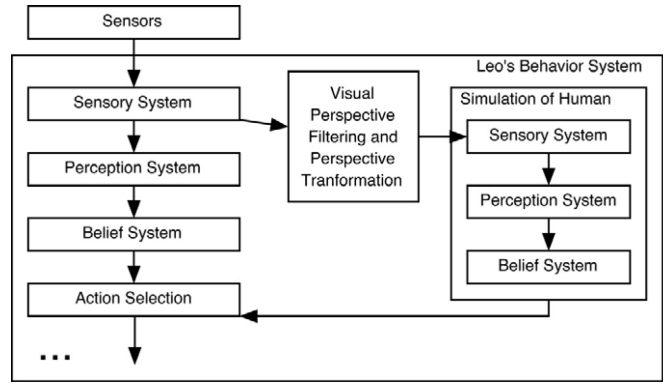


Fig. 2. Architecture for modeling the human's beliefs re-uses the robot's own architecture for belief maintenance.

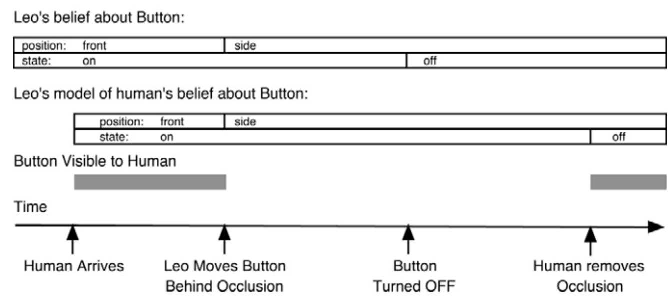


Fig. 3. Timeline following the progress of Leo's beliefs for one button. Leo updates his belief about the button with any sensor data available. However, Leo only integrates new data into his model of the human's belief for that button if the data is available when the human is able to perceive it.

and:

$$P : \{\text{robot local observations}\} \rightarrow \{\text{person local observations}\}. \tag{11}$$

Visibility can be determined by a cone calculated from the human's position and orientation, and objects on the opposite side of known occlusions from the human can be marked invisible.

Maintaining this parallel set of beliefs is different from simply flagging the robot's original beliefs as human visible or not, because it re-uses the entire architecture which has mechanisms for object permanence, history of properties, etc. This allows for a more sophisticated model of the human's beliefs.

Fig. 3 shows an example where this approach keeps track of the human's incorrect beliefs about objects that have changed state while out of the human's view. In this example, the human sees a button that is ON before it is moved behind an occluding barrier and turns OFF. Note that Leo's belief about the button is that it is OFF but the robot models the human's belief about the state of the button as still being ON. In this way, Leo has an initial set of mechanisms for modeling a human's potentially different beliefs. This technique has the advantage of keeping the model of the human's beliefs in the same format as the robot's own, allowing both for direct comparison between the two and for operating on these beliefs with the same mechanisms that operate on his own. This is important for establishing and maintaining mutual beliefs in time-varying situations where beliefs of individuals can diverge over time.

6. The task and goal learning mechanism

In our prior work in task and goal learning, the human interactively instructs the robot, building a new goal concept and task model representation from its set of known states, actions and tasks [3,12]. Each trial yields a number of potential hypotheses about the task and goal representation. Executing tasks and incorporating feedback narrows the hypothesis space, converging on the best representation. We have argued that flexible, goal-oriented, hierarchical task learning is imperative for learning in a collaborative setting from a human partner, due to the human's propensity to communicate in goal-oriented and intentional terms. In this work we extend the task and goal learning mechanism to learn from multiple perspectives.

We have a hierarchical, goal-oriented task representation. A task is represented by a set S of schema hypotheses: one primary hypothesis and n others. A schema hypothesis has x executables E (each either a primitive action a or another schema), a goal G, and a tally, c, of how many seen examples have been consistent with this hypothesis.

Goals for actions and schemas are a set of y *goal beliefs* about what must hold true in order to consider this schema or action achieved. A goal belief represents a desired change during the action or schema by grouping a belief's percepts into m criteria percepts and n expectation percepts. *Criteria percepts* indicate a feature that holds constant over the action or schema and *expectation percepts* indicate an expected feature change. This allows for a straightforward evaluation of an action or schema's goals during execution: for each goal belief, find all objects with the criteria features and check that the expectation features match.

Schema Representation:

$$S = \{[(E_1 \ldots E_x), G, c]_P, [(E_1 \ldots E_x), G, c]_{1\ldots n}\}$$

$$E = a \mid S$$

$$G = \{B_1 \ldots B_y\}$$

$$B = p_{C_1} \ldots p_{C_m} \cup p_{E_1} \ldots p_{E_n}.$$

Additionally, for the purpose of task learning, Leo can take a snapshot of the world (i.e. the state of the Belief System), $\mathrm{Snp}(t, x)$, in order to later reason about world state changes. The snapshot pertains to a time step t and can either be taken from Leo's (L) or the human's (H) belief perspectives, indicated by x.

Learning is mixed-initiative such that Leo pays attention to both his own and his partner's actions during a learning episode. When the learning process begins, Leo creates new schema representations S_{Leo} and S_{Hum} and saves belief snapshots $\mathrm{Snp}(t_0, L)$ and $\mathrm{Snp}(t_0, H)$.

From time t_0 until the human indicates that the task is finished (t_{end}), if either Leo or the human completes an action *act*, Leo makes an action representation, $a = [act, G]$, for both S_{Leo} and S_{Hum}:

For action *act* at time t_b given last action at t_a
G_L = belief changes from $\mathrm{Snp}(t_a, L)$ to $\mathrm{Snp}(t_b, L)$
G_H = belief changes from $\mathrm{Snp}(t_a, H)$ to $\mathrm{Snp}(t_b, H)$
append $[act, G_L]$ to exectuables of S_{Leo}
append $[act, G_H]$ to exectuables of S_{Hum}
$t_a = t_b$.

At time t_{end}, this same process works to infer goals for the schemas, S_{Leo} and S_{Hum}, making the goal inference from the differences in $\mathrm{Snp}(t_0, x)$ and $\mathrm{Snp}(t_{\mathrm{end}}, x)$. The goal inference mechanism notes all changes that occured over the task; however, there may still be ambiguity around which aspects of the state change are the goal (the change to an object, a class of objects, the whole world state, etc.). Our approach uses hypothesis testing coupled with human interaction to disambiguate the overall task goal over a few examples.

Once the human indicates that the current task is done, S_{Leo} and S_{Hum} contain the representation of the seen example $([(E_1 \ldots E_x), G, 1])$. Having been created from the same demonstration, the executables will be equivalent, but the goals may not be equal since they are from differing perspectives. In the case that they are different, Leo attempts to resolve this conflict by querying the human in order to choose a single schema representation S.

The system takes the goal from the human's schema, G_{Hum}, and creates the set *conflicts*, the parts of this goal that are incomplete from Leo's perspective: $\forall B \in G_{\mathrm{Hum}}; \forall b \in$ Leo's Belief System, if $\exists b$ s.t. b matches criteria percepts $p_{C_1} \ldots p_{C_m}$ in B, but does not match all expectation percepts, $p_{E_1} \ldots p_{E_n}$ in B, add b to conflicts.

Next, Leo makes a query to the human by making a pointing gesture towards the object represented by the first belief in conflicts. This query is meant to communicate, "Do we need to take care of this, too?" to the human partner. The system is persistent about the need to resolve the conflict and will continue to make this query to the human every 45 seconds until the human responds verbally with "Yes" or "No".

If the human makes a positive response, Leo assumes that $S = S_{\mathrm{Hum}}$ is the representation that the human intended to teach. Leo then expands the actions to resolve the set conflicts (i.e. $\forall b \in$ *conflicts* do all actions necessary to make expectation percepts match G_{Hum}), and completes this set of actions. Otherwise, if the human responds negatively, Leo assumes that $S = S_{\mathrm{Leo}}$ is the correct schema representation.

The system uses S to expand other hypotheses about the desired goal state to yield a hypothesis of all goal representations G consistent with the current demonstration (for details of this expansion process, see [12]). The current best schema candidate (the primary hypothesis) is chosen through a Bayesian likelihood method: $P(h|D) \propto P(D|h)P(h)$. The data, D, is the set of all examples seen for this task. $P(D|h)$ is the percentage of the examples in which the state change seen in the example is consistent with the goal representation in h. For priors, $P(h)$, our algorithm prefers a more specific hypothesis over a more general one (as determined by the number of goal beliefs, and number of criteria and expectation features in those beliefs). Thus, when a task is first learned, every hypothesis

Table 1
Social cues for scaffolding

Context	Leo's expression	Intention
Human points to object	Looks at Object	Shows Object of Attention
Executing an action	Looks at Object	Shows Object of Attention
Human: "Let's learn task X"	Subtle Head Nod	Confirms start of task X
Human: "Task X is done"	Subtle Head Nod	Confirms end of task X
Any speech	Perks ears	Conveys that Leo is listening
Unconfident task execution	Glances to human frequently	Conveys uncertainty
Completion of demonstration	Perks ears, lean forward	Soliciting feedback from teacher
Human: "Can you...?"	Perform or Nod/Shake	Communicates task knowledge
Human: "Do task X"	Performs X	Demonstrating hypothesis for X
Task done; Human: "Not quite..."	Subtle nod	Confirms feedback, expects refinement
Task done; Human: "Good!"	Nods head	Confirms task hypothesis

schema is equally represented in the data, and the algorithm chooses the most specific schema for the next execution.

7. Communication for social engagement during teaching scenarios

Human-style tutelage is a social and a collaborative process [9,16] and usually takes the form of a dialog, which is a fundamentally cooperative activity [11]. To be a good instructor, one must maintain an accurate mental model of the learner's state (e.g., what is understood so far, what remains confusing or unknown) in order to appropriately structure the learning task with timely feedback and guidance. The learner (robot or otherwise) helps the instructor by expressing its internal state via communicative acts (e.g., expressions, gestures, or vocalizations that reveal understanding, confusion, attention, etc.). Through reciprocal and tightly coupled interaction, the learner and instructor cooperate to help the instructor maintain a good mental model of the learner, and to help the learner leverage from instruction to build appropriate models, representations, and associations.

The robot cooperates in the teaching/learning collaboration by maintaining a mutual belief with the teacher about the task state, expressing confusion, understanding, attention, etc. A number of expressive skills contribute to Leo's effectiveness in learning through collaborative dialog (Table 1). Eye gaze establishes joint attention, reassuring the teacher that the robot is paying attention to the right thing. Subtle nods acknowledge task stages, confirming a mutual understanding of moving on to the next stage.

As Table 1 shows, we have given our robot a number of social and expressive skills that contribute to the robot's effectiveness in learning through collaborative discussion. For example, joint attention is established both on the object level and on the task structure level. The robot uses subtle expressions to indicate to the human tutor when it is ready to learn something new, and its performance of taught actions provides the tutor with immediate feedback about comprehension of the task. Envelope displays such as gaze aversion, eye contact and subtle nods are used to segment a complex task learning structure in a natural way for the tutor.

These social cues also help to "trouble-shoot" the interaction. For instance, if the robot is unable to parse the human instructor's utterance, Leo gestures by leaning forward with hand to ear to indicate that it failed to comprehend what was said, and to prompt the human to repeat their last phrase.

The robot's demonstration of the task being learned provides the human instructor with immediate feedback about the robot's current task comprehension. When demonstrating a task that it is currently trying to learn, the robot's ear pose, body position, and eye gaze are used to solicit feedback from the human when uncertainty is high. For instance, the schema hypothesis used for execution has a likelihood (between 0 and 1) relative to the other hypotheses available. If this confidence is low ($<.5$), Leo expresses tentativeness by frequently looking between the instructor and an action's object of attention to solicit feedback and further examples.

Upon finishing the task, Leo leans forward with his ears perked waiting for feedback. When the teacher confirms success through positive verbal feedback, Leo considers the task complete. Alternatively, if Leo has not yet achieved the goal, the instructor can give negative verbal feedback. Leo then expects the teacher to teach him the completion of the task. A new example is created through this refinement stage, similar to the original learning process.

8. Demonstration and discussion

In our experimental scenarios, a human instructor can teach Leonardo a variety of tasks involving a set of colored buttons [3,12]. The buttons can be pressed ON or OFF, switching an internal LED so that the robot can visually perceive if the button has been "activated" or "deactivated". Tasks involve learning specific patterns of button activation and deactivation, or learning concepts that represent new task goals (e.g., generalizing what it means to turn all the buttons on for any number of buttons in the workspace).

In this work, we explore two issues in the teaching scenario. First, the interaction consists of mixed-initiative demonstrations by both human teacher and robot learner. The human can demonstrate actions on objects for the robot to learn from, or verbally and gesturally direct the robot in what to do for each trial. Second, we explore socio-cognitive issues that arise in teaching the robot a new task. In particular, we explore the case where the perceived context is different from the human's perspective than from the robot's.

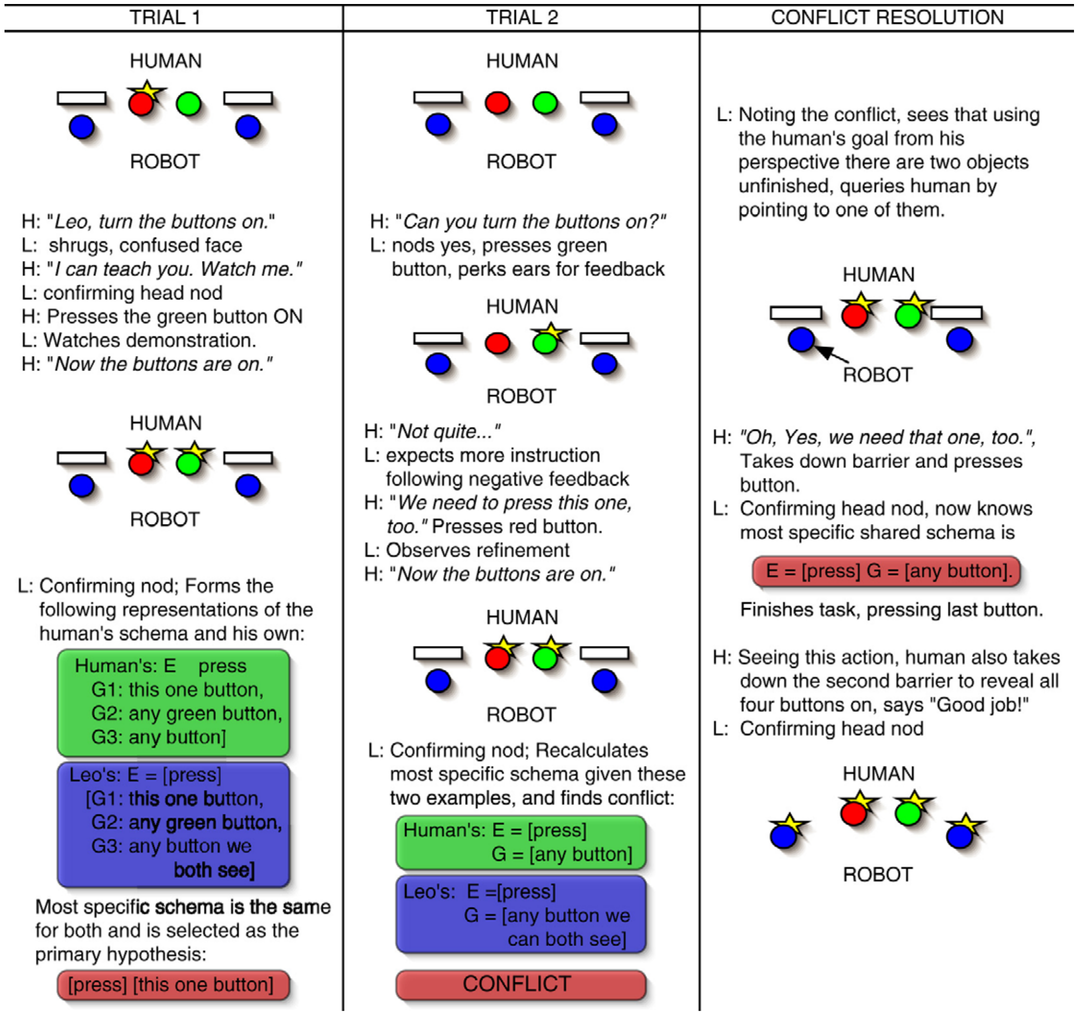


Fig. 4. Task Scenario. As Leo learns "turn all the buttons on", he considers the task from both his own and the human's perspective. Since the examples Leo has been given result in different task models depending on whether he considers the scenario from his perspective or the human's, he asks for clarification when requested to "turn all the buttons on" to determine which task model the human was intending. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

In general, how to handle and resolve "misconceptions" is a significant issue in developing collaborative systems capable of goal inference and plan recognition [5]. In this paper, we look at how misconceptions might arise in a collaborative teaching scenario, giving rise to "flawed" demonstrations. This causes the robot to entertain multiple, conflicting hypotheses in its attempt to infer what the human is trying to teach it to do. The robot must therefore be able to recognize when misconceptions are present and be able to collaborate with the teacher to resolve these problems in order to learn the intended task.

Fig. 4 depicts a scenario that demonstrates our mixed-initiative learning interaction. In this scenario, we start with four colored buttons (1 red, 2 blue, and 1 green) that the robot and human can press ON or OFF. The human and robot are positioned on opposite sides of the shared workspace. In addition, there are two occluding barriers in the workspace. Both blue buttons are occluded from the human (H), but not from Leonardo (L).

In this scenario the human teaches Leo the task: "turn all the buttons on". In the first trial, having asked if Leo knows how to do the task and received a negative answer, the human offers to teach Leo the task and then demonstrates the one action necessary to complete the task. The learning mechanism begins a learning process for two schemas, one based on Leo's perspective and one based on the human's perspective. Leo infers the goal for the schema based on the differences over the task. In both the human's and Leo's schema, the most specific schema for the task is to *press this green button ON*. In the human's schema, there is also some nonzero probability that the task is to *press any button ON*, but since Leo can see that the hidden blue buttons are OFF, he does not consider this hypothesis for himself. However, he does consider to *press any button we both see ON* as a possible schema for this example that accounts for the mutual evidence presented so far.

In the second trial, Leo uses his new schema representation to try and complete the task, pressing the green button ON. Leo looks back to the human, nodding to signal that he thinks the task is done, with ears perked to elicit feedback, since he is not confident he has learned the task (other hypotheses are being entertained). The human says, "Not quite", and Leo understands this as negative feedback and waits for further instruction or demonstration for refinement of his hypothesis. Leo watches the human's demonstration, recognizes the action, infers goals for this action based on both his beliefs and the human's, and modifies the schemas for this task, updating the most confident hypothesis, given this second example. Now, from the human's perspective, the most likely hypothesis schema is to *press any button ON*. But from Leo's perspective, this hypothesis does not match the evidence he has seen, so his best schema is *press the buttons we can both see ON*.

The learning mechanism then starts a process to resolve these two conflicting schemas. The process determines that if Leo were to use the human's schema in his current perspective, there are still two (hidden) buttons that remain undone. Leo acts to resolve this ambiguity by pointing to one of the hidden blue buttons and looking to the human with a questioning expression. The human takes down the barrier and confirms that these are part of the goal. Leo updates his hypothesis based on this observation to the shared hypothesis schema of *press any button ON*. He completes the task by pressing the remaining blue button.

This example is in a low-dimensional feature space and little data is needed to resolve the ambiguities, but even here the advantage of the tutelage paradigm is shown. Through a turn-taking interaction with a human partner, the robot quickly acquires the representative examples needed to generalize to the correct task representation. The robot's expressions and demonstrations of learned abilities help the teacher understand what the robot knows and what ambiguities remain. The robot is able to detect and resolve potentially ambiguous demonstrations, and resolves these collaboratively with the human to maintain mutual belief as to the generalized meaning of the intended concept to be learnt.

9. Conclusion

We have presented an architecture for a robot that learns simple tasks from flawed or ambiguous demonstrations by taking the perspective of the teacher. Assuming the perspective of the teacher allows the robot to guess the human's goals even if the human fails to properly achieve them.

In many cases where the robot and human have different knowledge about the world, the human could be attempting to teach the robot with incomplete information. In the example given here, the human thinks that she is teaching Leo the concept of "turn on all the buttons". However, since the human does not see all the buttons, the human is not correctly demonstrating the actions to the robot. Based on his perspective-taking abilities, Leo can model this error and understand the human's intent without any correct examples.

This same technique helps with many kinds of failed teaching scenarios. Imagine a human attempting to activate a push control, mistaking it for a twist control. If Leo could model this false knowledge, then he could determine that they meant to activate the control by pushing it even though he has never seen it happen.

We believe that perspective taking is an important part of creating an intuitive learning interaction with a human. Others have argued that it is critical for collaboration on a shared task in a physical space [6]. In our collaborative learning scenario, the robot has the advantage of having a human teacher who is attempting to teach the robot a particular skill as best they can with the information they have. It is important to make use of this situation to learn as much as possible. The architecture presented here takes advantage of the presence of the helpful human by learning in an interactive setting that allows the human to correct errors in the task model of the robot immediately as they happen, allowing the robot to more easily determine the specific task error and learn from few examples. Further, the robot's ability to take the perspective of the human to learn from their intent instead of their raw actions allows the robot to make use of well-intentioned but flawed or ambiguous examples.

Acknowledgements

The work presented in this paper is a result of ongoing efforts of the graduate and undergraduate students of the MIT Media Lab Robotic Life Group. This work is funded by the *Digital Life* and *Things That Think* consortia of the MIT Media Lab.

References

[1] L.W. Barsalou, P.M. Niedenthal, A. Barbey, J. Ruppert, Social embodiment, The Psychology of Learning and Motivation 43 (2003).

[2] B. Blumberg, M. Downie, Y. Ivanov, M. Berlin, M.P. Johnson, B. Tomlinson, Integrated learning for interactive synthetic characters, in: Proceedings of ACM SIGGRAPH 2002, ACM Transactions on Graphics 21 (3) (2002).

[3] C. Breazeal, G. Hoffman, A. Lockerd, Teaching and working with robots as collaboration, in: Proceedings of the AAMAS, 2004.

[4] C. Breazeal, C. Kidd, A. Lockerd Thomaz, G. Hoffman, M. Berlin, Effects of non verbal communication on efficiency and robustness in human-robot teamwork, in: IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, 2005.

[5] S. Carberry, Techniques for plan recognition, User Modeling and User-Adapted Interaction 11 (1–2) (2001) 31–48.

[6] N.L. Cassimatis, J.G. Trafton, M.D. Bugajska, A.C. Schultz, Integrating cognition, perception and action through mental simulation in robots, Journal of Robotics and Autonomous Systems 49 (1–2) (2004) 13–23.

[7] M. Davies, T. Stone, Introduction, in: M. Davies, T. Stone (Eds.), Folk Psychology: The Theory of Mind Debate, Blackwell, Cambridge, 1995.

[8] V. Gallese, A. Goldman, Mirror neurons and the simulation theory of mind-reading, Trends in Cognitive Sciences 2 (12) (1998) 493–501.

[9] J. Glidewell (Ed.), The Social Context of Learning and Development, Gardner Press, New York, 1977.

[10] J. Gray, C. Breazeal, M. Berlin, A. Brooks, J. Lieberman, Action parsing and goal inference using self as simulator, in: 14th IEEE International Workshop on Robot and Human Interactive Communication (ROMAN), Nashville, Tennessee, IEEE, 2005.

[11] H.P. Grice, Logic and conversation, in: P. Cole, J.L. Morgan (Eds.), Syntax and Semantics, in: Speech Acts, vol. 3, Academic Press, San Diego, CA, 1975, pp. 41–58.

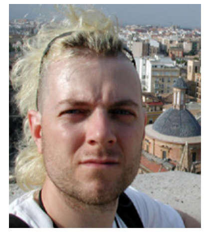
[12] A. Lockerd, C. Breazeal, Tutelage and socially guided robot learning, in: IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, 2004.

[13] A.N. Meltzoff, J. Decety, What imitation tells us about social cognition: a rapprochement between developmental psychology and cognitive neuroscience, Philosophical Transactions of the Royal Society: Biological Sciences 358 (2003) 491–500.

[14] M.N. Nicolescu, M.J. Matarić, Natural methods for robot task learning: Instructive demonstrations, generalization and practice, in: Proceedings of the Second International Joint Conference on Autonomous Agents and Multi-Agent Systems, Melbourne, Australia, July 2003.
[15] S. Schaal, Is imitation learning the route to humanoid robots? Trends in Cognitive Sciences 3 (1999) 233–242.
[16] L.S. Vygotsky, Mind in Society: The Development of Higher Psychological Processes, Harvard University Press, Cambridge, MA, 1978.

Cynthia Breazeal is an Associate Professor of Media Arts and Sciences at the MIT Media Lab, and is Director of the Robotic Life Group, holding the LG Group career development chair. Her research focuses in the area of Human-Robot Interaction where the robot is designed to interact with people as a socially intelligent partner. Her current research explores human-robot teamwork and collaboration, social cognition, and social forms of learning such as imitation and tutelage. Cynthia Breazeal received her S.M (1993) and Sc.D. (2000) in Electrical Engineering and Computer Science from the Massachusetts Institute of Technology at the Artificial Intelligence Lab.

Matt Berlin received his A.B. in Computer Science from Harvard College in 2001, and his S.M. in Media Arts and Sciences from MIT in 2003. His research focuses on applying models of learning and development from ethology and cognitive psychology to the design of robots and animated characters. He is currently a doctoral student at the MIT Media Lab, developing robots that can engage in sensible, structured learning interactions with human teachers.



Andrew Zoz Brooks received his undergraduate education at the University of Adelaide in South Australia, majoring in Organic Chemistry and Computer Science and received First Class Honors in Computer Science and a Graduate Diploma in Education. He received his Master's degree at the Robotic Systems Laboratory of the Research School of Information Sciences and Engineering at the Australian National University in Canberra. He is currently a doctoral student at the MIT Media Lab, with current research interests including robot learning from demonstration and playful human–robot interaction.

Jesse Gray received an undergraduate degree in mathematics and computer science from Tufts University, and an S.M. at the MIT Media Lab, where he is currently a Ph.D. student in the Robotic Life Group. Jesse's recent work has been looking at applying concepts from simulation theory, borrowed from cognitive psychology, to help robots understand observed behavior in terms of the robot's own behavior generation mechanisms.

Andrea L. Thomaz received her B.S. in Electrical and Computer Engineering from the University of Texas at Austin in 1999, and her S.M. in Media Arts and Sciences from the Massachusetts Institute of Technology in 2002, where she is currently a Ph.D. candidate at the Media Lab working in the Robotic Life Group. Her research centers on building machines with social intelligence, focused currently on socially guided learning, exploring the ways in which machine learning can be designed to more fully take advantage of human interaction and tutelage.