Optimality of Human Teachers for Robot Learners

Maya Cakmak and Andrea L. Thomaz

Center for Robotics and Intelligent Machines, Georgia Institute of Technology

801 Atlantic Dr., Atlanta GA 30332

{maya, athomaz}@cc.gatech.edu

Abstract—In this paper we address the question of how closely everyday human teachers match a theoretically optimal teacher. We present two experiments in which subjects teach a concept to our robot in a supervised fashion. In the first experiment we give subjects no instructions on teaching and observe how they teach naturally as compared to an optimal strategy. We find that people are suboptimal in several dimensions. In the second experiment we try to elicit the optimal teaching strategy. People can teach much faster using the optimal teaching strategy, however certain parts of the strategy are more intuitive than others.

I. Introduction

Our research aims at building social robots that can learn new skills and tasks from everyday people. Many learning problems in this scenario fit into the Supervised Learning paradigm where a teacher provides examples of members and non-members of a concept, and the learner builds models from these examples. How well and how fast a new concept is learned depends both on the learning algorithm and the data provided. While humans may have a good intuition about what examples to give to a robot learner, they might not be complete or optimal since they lack a good understanding of the underlying learning algorithm. Our approach, *Socially Guided Machine Learning* (SGML), advocates taking advantage of the ways people naturally teach as well as trying to improve the way people teach in order to better match the learner.

There is considerable research related to agents that learn interactively from humans [Chernova and Veloso, 2008], [Grollman and Jenkins, 2008], [Thomaz and Breazeal, 2008]. These have started to take a SGML approach, giving more consideration to learning from everyday people. However the focus has been on trying to improve the learner, not the teacher.

The teaching scenario has also been addressed in Machine Learning Theory. Machine learning algorithms are often designed assuming that data comes from an average teacher (*i.e.* a fixed but unknown distribution) or an adversarial teacher who tries to make learning as difficult as possible. The goal is to build learning algorithms that can learn from *any* teacher. However the output of a learning algorithm is often highly dependent on the data set, and thus dependent on the teacher. A good teacher can produce a data set that results in a more accurate output with less examples. Producing such a data set for an arbitrary concept is a challenging problem and has been formally studied within the field of *algorithmic teaching* [Balbach and Zeugmann, 2009], [Mathias, 1997], [Goldman and Kearns, 1995].

Identifying lower bounds on the complexity of teaching a concept requires coming up with an algorithm that can teach

Fig. 1. The Simon robot interacting with a teacher.

the concept with minimum number of examples. In certain cases this number is much smaller then the average number of examples used for learning from randomly chosen examples. Furthermore, even an active learner that can choose which examples to learn from can never learn faster than when a passive learner is taught optimally [Angluin, 1988], [Goldman and Kearns, 1995]. Therefore having a helpful teacher can considerably improve the learning rate.

The question we address in this paper is how closely everyday human teachers match a theoretically optimal teacher. We first collect data in an experiment where people teach a concept to our robot in a supervised fashion (giving a series of labeled examples) with no instruction on teaching strategy. We find that people are able to teach the robot successfully, but not optimally. A second experiment investigates whether or not we can give people simple prompts to elicit the optimal teaching strategy. In one condition we simply tell people to teach with as few examples as possible, and in the second condition we explain the optimal strategy given in [Goldman and Kearns, 1995]. We find that people can effectively use the optimal strategy to teach much faster, but some aspects of the optimal teaching strategy are more intuitive than others. These results inform the design of robot learners, pointing out that a good place for the learner to intervene in the learning process is in helping the human select optimal negative examples.

II. Approach

A. Robot Platform

The robot platform used in this study is "Simon," an upper-torso humanoid 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 with a human partner at a tabletop workspace.

B. Domain Description

In this work, Simon's learning task involves colorful paper cutouts, which we call *tangrams*. Each tangram has a *color* (pink, green, yellow, or orange), a *shape* (square, triangle, or circle), and a *size* (small or large) feature, for a total of 24 possible unique tangrams. Simon learns concepts from labeled examples of *compound tangrams*—two piece configurations. A compound tangram has 6 features: $color_{top}$, $shape_{top}$, $size_{top}$, $color_{bottom}$, $shape_{bottom}$, and $size_{bottom}$. Simon's workspace contains exactly one of each tangram piece, so there are 552 (24×23) possible compound tangrams.

Simon's workspace is a table covered by a black tablecloth, with the 24 tangrams pieces arranged on the perimeter of the table. 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 tangrams in the demonstration area when the teacher provides a label. Tangrams are detected through a fixed overhead camera and segmented using background subtraction. The shape of the tangram is recognized by the number of corners of the simplified polygon contour of the segmented object (square: 4, circle: 8, triangle: 3). Size is based on the area within the contour and color is recognized using the color histogram of the segmented object. The tangrams are localized in robot world coordinates using a fixed homography from the image plane to the table plane.

Labels or test questions are provided by the teacher using a fixed set of sentences. These sentences are recognized using the Sphinx speech recognition system. In order to reduce errors, both vision and speech perception are monitored and corrected by an experimenter during the interaction.

D. Actions

Simon uses speech synthesis and gaze directions to interact with the human teacher. He looks up and blinks its ears to indicate that it is ready for another example. To acknowledge that the example given by the teacher has been processed, Simon gives a verbal confirmation such as "okay" or "thank you." Simon uses a gesture (head nod or shake) in conjunction with speech ("Yes, this is a house", "No, it's not" or "I don't know") to respond to tests. He randomly chooses a sentence from a fixed set of sentences that have the same meaning.

E. Interaction

The learning/teaching process involves a turn-taking interaction between Simon and the human teacher. First the teacher prepares a compound tangram instance in the demonstration area and labels it or asks Simon a test question. When the

TABLE I
CONCEPTS

Name	Concept/Hypothesis (h^*)	Examples	# of Instances
HOUSE	$shape_{top} = triangle \wedge$ $color_{top} = pink \wedge$ $shape_{bottom} = square$		16
ICE CREAM	$shape_{top} = circle \wedge$ $shape_{bottom} = triangle \wedge$ $color_{bottom} = yellow$		16

sentence is heard, Simon gazes towards the demonstration area and perceives the instance. If the person labels the instance, Simon incorporates this new labeled example into the concept model and then acknowledges the example. If the person asks a test question, Simon classifies the instance based on the current model of the concept and responds to the test. The teacher's next turn starts when Simon blinks his ears.

III. LEARNING AND TEACHING

This section describes the learning and teaching problems considered in this study. We note the fundamental difference betweens learning and teaching as follows. A learning algorithm produces a hypothesis in some concept class based on a set of labeled examples consistent with an *unknown* target concept h^*. A teaching algorithm produces a set of labeled examples based on a *known* concept h^*. In this study the robot is the learner and the human subject is the teacher.

A. The Concept Class

The *concept class* used in this study is the set of monotone conjunctions of discrete-valued variables (features). The variables are not necessarily binary since they can have more than two discrete values (*e.g.* color) and there are at most six variables in the conjunction. Therefore, a *concept* or *hypothesis* is defined by a conjunction of compound tangram feature values that must hold true to be a member of that concept. An example concept is HOUSE, as described in Table I. A HOUSE is a compound tangram that has a pink and triangular top piece and a square bottom piece. The size of either piece and the color of the bottom piece do not matter. Thus, the concept HOUSE is represented with the conjunction $\{color_{top} = pink \wedge shape_{top} = triangle \wedge shape_{bottom} = square\}$.[1] The features that appear in the conjunction of the target concept are referred to as *relevant* features. In this experiment, Simon learns the concepts given in Table I.

B. Concept Learning

For concept learning we use a variant of the halving algorithm [Littlestone, 1988]. The algorithm maintains the subset of all possible hypotheses which contains hypotheses that are most consistent with the given examples. This subset is called

---

[1] Assuming the order of features given in Sec. II-B, we use the shorthand: $< pink, triangle, *, *, square, * >$ where $*$ means the value does not matter.

TABLE II
PROGRESS OF THE VERSION SPACE AS EXAMPLES ARE PROVIDED.

Step	Example	Label	Version Space (V)
1		+	**Most specific hypothesis:** $<pink,triangle,large,yellow,square,large>$ **Most general hypotheses:** $<pink,*,*,*,*,*>$, $<*,triangle,*,*,*,*>$ $\ldots <*,*,*,*,*,large>$
2		−	**Most specific hypothesis:** $<pink,triangle,large,yellow,square,large>$ **Most general hypothesis:** $<pink,*,*,*,*,*>$
3		+	**Most specific hypothesis:** $<pink,triangle,*,yellow,square,*>$ **Most general hypothesis:** $<pink,*,*,*,*,*>$

Fig. 2. An optimal teaching sequence example for target concept HOUSE.

the *version space* (V). With each new example, the algorithm updates the consistency count of all hypotheses and prunes V. An example is shown in Table II. Using this algorithm the robot can make online predictions while learning, based on the majority in the predictions of the hypotheses in V. If the disagreement in the predictions is high, the robot shows uncertainty in its response to a test (by saying "I don't know"). The learning algorithm is said to have *exactly identified* the target concept when V includes a single hypothesis which is the target concept (h^*).

C. Concept Teaching

A traditional way to measure complexity of teaching a concept class is the *teaching dimension* [Goldman and Kearns, 1995]. This is the minimum number of labeled examples needed to uniquely identify *any* concept in a concept class. A sequence of examples that uniquely identifies a target concept is called a *teaching sequence* denoted by T. The teaching dimension of the concept class in this study is identified as $min(r+2, n+1)$ where r is the number of relevant variables (features) and n is the total number of variables. For the two concepts in our study $min(|T|)$=5 since r=3 and n=6.

The optimal teaching strategy to achieve this lower bound is as follows [Goldman and Kearns, 1995]. The teacher first gives a positive example and then another positive example where all the $n-r$ irrelevant features are changed to a different value. This demonstrates that the features that changed between the two examples are irrelevant. Next the teacher needs to prove that the features that were not changed are actually relevant. This is achieved by changing the relevant features one at a time and label the instance as negative. Since only one feature is different from what is known to be a positive example, the change proves that the feature is a relevant one. An example optimal teaching sequence for HOUSE is shown in Fig. 2.

Natarajan, 1989 uses a different measure of teaching complexity which is the number of examples that must be given until the most specific hypothesis consistent with the data is is the target concept. In the strategy described above, this criteria is met after the first two positive example, therefore the complexity according to this measure is 2.
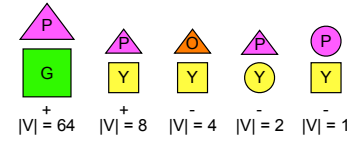
IV. EXPERIMENTS

We performed two experiments to investigate how people naturally teach concepts and whether human teaching can be made more optimal. In the first experiment people taught a concept to our robot Simon through a face-to-face interaction. In the second experiment we asked people to teach two concepts to a virtual Simon and prompted them before the second concept to use a particular teaching strategy. In this section we describe the details of these two experiments and present the metrics we use to measure the teaching behavior.

A. Experiment 1

In the first experiment 24 subjects taught Simon the HOUSE concept. In order to teach Simon, subjects arrange a compound tangram in the demonstration area from Simon's perspective and say one of three sentences:

- [Simon], this is a HOUSE. (Positive label)
- [Simon], this is *not* a HOUSE. (Negative label)
- [Simon], is this a HOUSE? (Test)

Simon processes this statement and responds to it. Subjects were instructed to wait for Simon to blink the lights on his ears before continuing. They were told to continue teaching the concept until they were satisfied that Simon had learned it well or thought that he had stopped making progress.

B. Experiment 2

In the second experiment 20 subjects taught two symbols (HOUSE and ICE-CREAM) to a simulated robot using the interface shown in Fig. 3. In this interface, the tangrams are dragged to the demonstration area using the mouse. The instance in the demonstration area is labeled or tested using the buttons at the top which correspond to the three sentences used in Experiment 1. Subjects were told to use the "I'm done" button when they thought Simon had learned the concept.

The description of the concepts were given to the subject before starting to teach and a reminder was shown on the right of the interface. All subjects taught the first concept without any specific teaching instructions. Then subjects were prompted with specific instructions before teaching the second symbol. The two experiment groups had different prompts:

Group 1 (Motivated): "This time try to teach Simon with as few examples and questions as possible."

Group 2 (Strategy): "This time we want you to use a specific teaching strategy: (1) First, show Simon an example of ice-cream. (2) Then, change everything that doesn't matter for ice-cream, and show another ice-cream. (3) Then show examples of objects that are not ice-creams, but that differ from ice-cream by only one property. Change the properties that matter one-by-one."
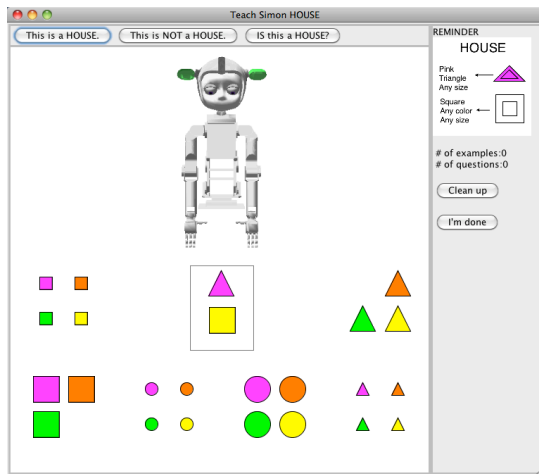
Fig. 3. Graphical interface used for teaching Simon in Experiment 2.

The instructions were repeated to the subjects by the experimenter and their questions were answered. A reminder of the instructions was also shown on the right part of the interface.

C. Metrics of Good Teaching

In this section we define a set of metrics to characterize how well people teach compared to an optimal teaching strategy.

1) Quality of teaching: Optimal teaching requires exact identification, however human teachers can often stop teaching without uniquely identifying the target concept. We measure how close someone gets to exactly identifying a concept with the number of hypotheses that remain in the version space after the example sequence is given (denoted by $|V|$).

Note that the learner can achieve perfect prediction performance without exact identification[2]. To capture the prediction performance of the learner we use the F-measure of the V produced by the learner:

$$F = 2 * \frac{precision * recall}{precision + recall}$$

In our experiments we measure how well a concept is taught with both $|V|$ and F values. We also look at the percentage of subjects that achieve $|V|=1$ and $F=1$ at the end of their example sequence (denoted by $\%(|V|=1)$ and $\%(F=1)$).

2) Speed of teaching: Optimal teaching requires the teacher to give as few examples as possible. If a teacher is able to exactly identify a concept their speed of teaching can be measured as the length of their teaching sequence, denoted by $|T|$. Similarly, the number of examples to get to $F=1$ is a measure of how fast one teaches. We denote this by $|T|_F$.

Since most teaching sessions do not result in exact identification of the target, a better measure is the rate at which the version space is pruned or the rate at which the learner's prediction performance improves. To capture this we use the average $|V|$ and the average F-score over the first few examples (denoted by $avg(|V|)$ and $avg(F)$). This is equivalent to

[2]This happens when V contains a set of hypotheses where all but one are split 50/50 in their prediction, and the last hypothesis is h^*.

measuring the area under the learning curve. We look at the first 5 examples since $min(|T|) = 5$.

A measure of the speed of teaching can include or exclude the tests performed by the teacher between labeled examples. Ideally a teacher asks no questions, however in practice human teachers require tests to verify what the learner knows because they might not have an accurate mental model of the learner or they might forget what they have already taught. An efficient teacher is expected to ask as few questions as possible. We investigate both the *number of examples* (excluding tests) and *number of steps* (including tests) to identify how fast people teach. The inclusion of *tests* in the teaching sequence or the averaging is denoted by the subscript t (T_t or avg_t). In addition to number of examples, we measure average wall-clock time taken to provide an example or test and we analyze the total number of tests performed while teaching.

3) Similarity to optimal teaching strategy: In order to better characterize the optimality of how humans teach we try to describe the "goodness" of examples given by the person according to the optimal teaching strategy given in Sec. III-C. We identify three tricks that this strategy uses in order to achieve accurate and fast learning:

The first trick is to start with a positive example, drastically pruning the hypothesis space (from all possible concepts to all that are consistent with the example, *i.e.* from 576 to 64).

The second trick is to vary as many irrelevant features as possible while keeping the second example positive. This has the potential to reduce the size of the version space by a factor of $2^{(n-r)}$ where $(n-r)$ is the number of irrelevant features. Thus, in our case a very "good" positive example reduces the size of the version space by a factor of $2^3 = 8$. Positive examples that vary by 2 or 1 irrelevant features will be less effective, with a pruning factor of 4 or 2, but still informative. If a positive example does not prune the hypothesis space at all it is a redundant, therefore "bad" positive example.

The final trick is to vary relevant features one at a time. This results in pruning half of the version space. When two or more features are varied at once, it is not possible to judge which subset of the features that are varied causes the instance to be a negative example. Therefore, the learner can only eliminate the hypotheses that have all of the varied features. We consider negative examples that halve the version space as "good" ones. The ones that do not change the version space (*i.e.* uninformative) or that change the version space by a factor less than 2 (*i.e.* too varied) are "bad" negative examples.

V. RESULTS

A. How optimal is natural teaching?

The goal of Experiment 1 was to analyze how subjects teach a concept when they are not given any specific instructions about how to teach. We analyze the data for teaching the HOUSE concept based on the teaching metrics. Table III presents the value of these metrics for both experiments together with optimal values.

1) Positive examples more optimal than negative: Human teachers seem naturally inclined to start with a positive example when teaching a concept; 91.67% (22/24) of subjects

started teaching with a positive example. Additionally, looking at the distribution of examples, we observe that positive examples given by teachers are mostly informative (more good examples than bad). Thus, people are able to induce variance in the irrelevant part of the space as part their effort to give "different" positive examples. In addition, the ratio of positive and negative examples in the data provided by people is very close to the ratio in an optimal teaching sequence (40% positive and 60% negative, Fig. 2).

2) Failed to achieve exact identification: The majority of the subjects ended teaching without achieving exact identification. A higher percentage reached an F-measure of 1, however there were still some subjects that ended teaching when the robot could not correctly classify some instances. When very few hypotheses are left in the version space, disagreement among the hypotheses occurs only for few instances and the majority prediction is correct for most instances. As a result it is difficult for subject to identify the example that will further prune the version space.

3) Speed of teaching is suboptimal: Subjects used twice the number of optimal examples to achieve $|V|$=1 and three times the number of optimal examples to achieve F=1. When including tests, the teaching sequence has 3-5 times more examples and tests than the optimal teaching sequence. The rate at which subjects teach at the beginning of the sequence is also suboptimal. For instance, the average F-measure throughout the first five examples is about half of the optimal value, which means that the robot will make many more mistakes if it is tested during this phase of teaching.

4) Example distribution is suboptimal: The underlying reasons for the suboptimal performance of natural teaching is better characterized by the distribution of good and bad examples (see Fig. 4(a)). We observe that very few subjects gave a positive example that prunes the version space by a factor of 8 (only 2 out of 24). Thus people do not seem to intuitively vary as much of the irrelevant space as possible. Second, we observe that only about one third of negative examples given by people is maximally informative.

5) Results confirmed in follow-up: The above findings were confirmed in the second experiment in which subjects were first asked to teach the HOUSE concept without any instructions on teaching strategy (see Experiment 2 baseline data in Table III and Fig 4(b)). The only difference was that the final performance of the robot (both $|V|$ and F) was closer to optimal in the Experiment 2 baseline. The faster pace of the simulation environment interaction makes it easier to perform many tests on the robot to determine what it has not learned yet, yielding better final performance.

B. Can we get people to be more optimal?

The goal of Experiment 2 was to characterize the ways in which people's teaching performance is altered when (i) they are motivated to be more efficient or (ii) they are asked to use a specific (optimal) teaching strategy. We analyze the data for teaching the ICE-CREAM concept in the two groups (Motivated and Strategy). Results are summarized in Table III and Fig. 4(c) and 4(d).
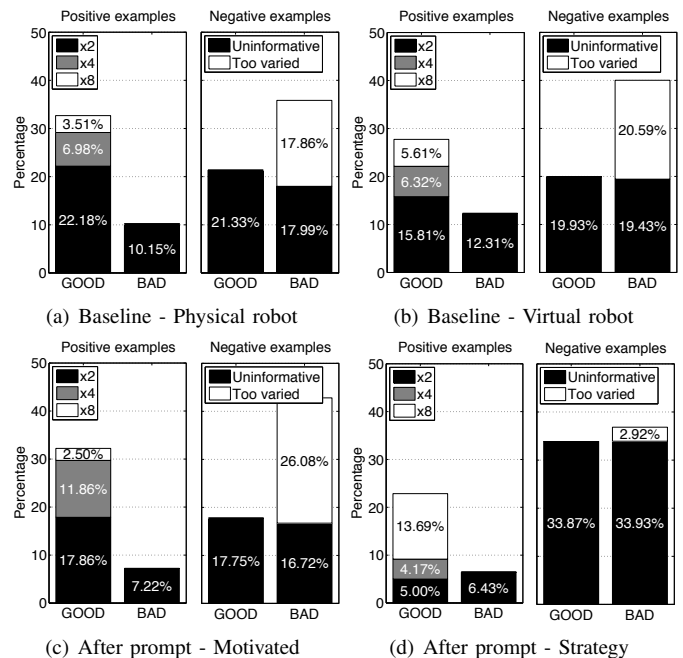

Fig. 4. Distribution of *good* and *bad* positive/negative examples. For good examples, we show breakdowns of how much of the version space was pruned (by a factor of 2, 4, or 8). For bad negative examples we breakdown whether this was due to the example being uninformative or too varied.

1) Optimal strategy not seen with generic prompt: The distribution of good and bad examples does not change considerably in the Motivated group (Fig. 4(c)) as compared to the Baseline groups (Fig. 4(a) and 4(b)). This shows that teachers cannot easily come up with a more optimal strategy on their own even when motivated to be more efficient.

2) Optimal strategy seems hard to grasp: When successfully adopted, the optimal strategy lets the teacher exactly identify the target concept with 5 examples. But in the Strategy group we saw that only 1 person out of 10 could teach in 5 examples. Most subjects were not able to fully adopt the strategy. People usually failed in giving the correct set of negative examples when following the strategy. While their negative examples were correctly chosen to have only one relevant feature different from a positive example, they failed at varying the feature that was different. Instead they varied the value of the feature that was different resulting in uninformative examples. As a result, the Strategy group has a high number of uninformative negative examples (Fig. 4(d)).

3) Quality of teaching: Compared to the baseline, fewer people achieve exact identification ($|V|$=1) after the prompt in both groups. We also see that the final $|V|$ and F are worse than the baseline after the prompt. We believe the reason is that in the Motivated group subjects stop teaching too early, trying to be efficient and in the Strategy group most fail to correctly follow the strategy.

4) Speed of teaching: While exact identification is not achieved, we observe that the average performance in the first 5 steps of teaching ($avg(|V|)$, $avg(F)$, $avg_t(|V|)$, $avg_t(F)$) is better for both groups after the prompt. Teaching in the Strategy group is significantly faster than the Motivated group.

We observe that the average time taken for providing

TABLE III
SUMMARY OF ALL TEACHING PERFORMANCE METRICS

Metric		Experiment 1	Experiment 2				Optimal Value
		Baseline Physical Robot	Baseline Virtual Robot	After prompt Motivated	After prompt Strategy	Comparison Motivated vs. Strategy	
Quality of Teaching	$\%(|V|=1)$	25% (6/24)	40% (8/20)	20% (2/10)	30% (3/10)	N/A	N/A
	$\%(F=1)$	58.33% (14/24)	90% (18/20)	70% (7/10)	90% (9/10)	N/A	N/A
	$|V|$	6.62 (S=11.79)	2.05 (SD=1.23)	8.10 (SD=11.68)	2.00 (SD=0.94)	t(18)=1.65, p>.05	1
	F	0.78 (SD=0.31)	0.97 (SD=0.10)	0.82 (SD=0.32)	0.94 (SD=0.19)	t(18)=-1.03, p>.05	1
Speed of Teaching (w/o tests)	$|T|$	10.67 (SD=4.59)	9.88 (SD=4.02)	6.00 (SD=0.00)	7.00 (SD=1.73)	t(3)=-0.77, p>.05	5
	$|T|_F$	6.93 (SD= 4.43)	6.00 (SD=2.22)	4.86 (SD=1.95)	3.00 (SD=2.12)	t(14)=1.80, p>.05	2
	$avg(|V|)$	48.87 (SD=68.56)	41.27 (SD=32.80)	26.22 (SD=9.43)	18.36 (SD=3.52)	t(18)=2.47, p<.05*	15.80
	$avg(F)$	0.36 (SD=0.19)	0.42 (SD=0.20)	0.52 (SD=0.20)	0.72 (SD=0.17)	t(18)=-2.41, p<.05*	0.82
Speed of Teaching (w tests)	$|T_t|$	19.17 (SD=7.41)	16.38 (SD=8.14)	19.50 (SD=19.09)	7.67 (SD=2.52)	t(3)=1.16, p>.05	5
	$|T_t|_F$	10.86 (SD=6.99)	8.28 (SD=6.29)	7.29 (SD=4.57)	3.11 (SD=2.42)	t(14)=2.36, p<.05*	2
	$avg_t(|V|)$	73.90 (SD=125.80)	43.84 (SD=32.96)	29.72 (SD=11.41)	18.44 (SD=3.51)	t(18)=2.99, p<.01*	15.80
	$avg_t(F)$	0.25 (SD=0.21)	0.37 (SD=0.25)	0.45 (SD=0.26)	0.72 (SD=0.17)	t(18)=-2.79, p<.05*	0.82

examples or tests (t_{avg}) is larger when subjects are prompted (Baseline: 10.43*sec*, Motivated: 11.01*sec*, Strategy: 18.06*sec*). In the Strategy group teachers were slower because they had to repeatedly go over the strategy and produce examples accordingly. In the Motivated group they spent more time considering the informativeness of each example to be more efficient. In addition, for both groups there is the extra mental load of trying to remember what they have already shown.

Even though subjects were slower in providing examples in the Strategy group, they needed less examples and they needed to ask less questions. As a result we observe that the expected total time ($t_{avg} * |T_t|_F$) is much smaller in the Strategy group (Baseline: 86.36*sec*, Motivated: 80.26*sec*, Strategy: 56.16*sec*).

5) Similarity to optimal example distribution: The percentage of highly informative positive examples (x8) is much higher in the Strategy group (Fig. 4(d)). Most subjects were able to successfully adopt the strategy for positive examples (7 of 10) and were able to teach the irrelevant features with fewer positive examples. As a result, the number of examples and tests required to achieve F=1 is significantly smaller in the Strategy group than the Motivated group. This is because the optimal strategy reaches F=1 after the first 2 positive examples and most subjects were correctly following this part of the strategy. Additionally, the Strategy group used less tests on average (Baseline: 9.10, Motivated: 7.90, Strategy: 6.10).

The number of uninformative positive examples went down in both groups (Fig. 4(c) and 4(d)). The percentage of negative examples with little information is smaller in the Strategy group (Fig. 4(d)). Subjects understood that negative examples should have small variance from positive examples to be maximally informative.

VI. CONCLUSIONS

In this paper we address the question of how closely everyday human teachers match a theoretically optimal teacher. We present two experiments in which subjects teach a concept to our robot in a supervised fashion. We find that people are suboptimal in several dimensions when they teach without any instructions, but they can get much faster when they know the optimal teaching strategy.

We observed that parts of the optimal strategy were less intuitive and we believe this points to an opportunity for queries from the robot that can help finish exact identification of the concept, demanding less from the teacher. In particular, people intuitively gave optimal positive examples, but had trouble with optimal negative examples. Thus, an ideal collaboration would only require the teacher to optimally teach the positive space and let the learner query the teacher to further identify the concept. Moreover, we would only expect the difficulties that human teachers had following the optimal strategy to be magnified in a more complex or realistic task domain. This gives further argument for an active robot learner that helps shape the human's input.

Even though our experiments involve a simple concept class and a particular learning algorithm, teaching strategies that significantly speed up learning might exist for more complicated concepts and learning algorithms. For example a generic rule for training classifiers could be to present a large variance within members of a class (*i.e.* few good representatives that span the whole class) and small variance between members and non-members of the class (*i.e.* borderline examples that reveal the limits of the class).

REFERENCES

[Angluin, 1988] Angluin, D. (1988). Queries and concept learning. *Machine Learning*, 2:319–342.

[Balbach and Zeugmann, 2009] Balbach, F. and Zeugmann, T. (2009). Recent developments in algorithmic teaching. In *3rd Intl. Conference on Language and Automata Theory and Applications*, pages 1–18.

[Chernova and Veloso, 2008] Chernova, S. and Veloso, M. (2008). Multithresholded approach to demonstration selection for interactive robot learning. In *3rd ACM/IEEE Intl. Conference on Human-Robot Interaction*.

[Goldman and Kearns, 1995] Goldman, S. and Kearns, M. (1995). On the complexity of teaching. *Computer and System Sciences*, 50(1):20–31.

[Grollman and Jenkins, 2008] Grollman, D. H. and Jenkins, O. C. (2008). Sparse incremental learning for interactive robot control policy estimation. In *IEEE International Conference on Robotics and Automation*.

[Littlestone, 1988] Littlestone, N. (1988). Learning quickly when irrelevant attributes abound: a new linear-threshold algorithm. *Machine Learning*, 2:285–318.

[Mathias, 1997] Mathias, H. D. (1997). A model of interactive teaching. *Computer and System Sciences*, 54(3):487–501.

[Natarajan, 1989] Natarajan, B. (1989). On learning boolean functions. In *19th Annual ACM Symp. on Theory of Computing*, pages 296–304.

[Thomaz and Breazeal, 2008] Thomaz, A. L. and Breazeal, C. (2008). Teachable robots: Understanding human teaching behavior to build more effective robot learners. *Artificial Intelligence Journal*, 172:716–737.