Opportunistic Active Learning for Grounding Natural Language Descriptions

Jesse Thomason[1*] **Aishwarya Padmakumar**[1*] **Jivko Sinapov**[2]

Justin Hart[1] **Peter Stone**[1] **Raymond J. Mooney**[1]

[1]Computer Science Department
University of Texas at Austin
{jesse, aish, hart, pstone, mooney}@cs.utexas.edu

[2]Computer Science Department
Tufts University
jsinapov@cs.tufts.edu

Abstract: Active learning identifies data points from a pool of unlabeled examples whose labels, if made available, are most likely to improve the predictions of a supervised model. Most research on active learning assumes that an agent has access to the entire pool of unlabeled data and can ask for labels of any data points during an initial training phase. However, when incorporated in a larger task, an agent may only be able to query some subset of the unlabeled pool. An agent can also opportunistically query for labels that may be useful in the future, even if they are not immediately relevant. In this paper, we demonstrate that this type of opportunistic active learning can improve performance in grounding natural language descriptions of everyday objects—an important skill for home and office robots. We find, with a real robot in an object identification setting, that inquisitive behavior—asking users important questions about the meanings of words that may be off-topic for the current dialog—leads to identifying the correct object more often over time.

Keywords: Grounded Language Learning, Active Learning, Human-Robot Interaction

1 Introduction

In machine learning tasks where obtaining labeled examples can be costly, active learning allows a system to select its own training data to obtain better performance using fewer labeled examples [1]. Active learning allows an agent to iteratively query for labels of examples from an unlabeled pool, selecting examples believed to be most useful for improving its model.

An important skill required by robots in a home or office setting is retrieving objects based on natural language descriptions. We consider an object retrieval task where humans can describe real-world objects using both visual and non-visual words (e.g. "red" and "heavy"). In this task, the pool of pre-labeled examples can be extremely limited, since curating a set of all words that apply to every object in an environment is a huge annotation effort for a human user, motivating the use of active learning to query for additional labels.

A robot in operation will typically be restricted to querying about objects that are physically nearby. In addition, it may be engaged in a task with the human to whom the query is addressed, to which the query may be unrelated. In such situations, the robot needs to be inquisitive—asking questions that may not be immediately relevant to the task at hand, *and* opportunistic—asking locally "convenient" questions that may not be optimal among all objects since only a subset of objects is available.

We call this setting *opportunistic active learning*, and it differs from existing work on active learning in three key ways. First, the agent may not be able to ask queries that are globally most useful

[*]These two authors contributed equally.

1st Conference on Robot Learning (CoRL 2017), Mountain View, United States.

to improve its models, since the task setting limits the available objects. Second, the agent must decide whether or not to ask such queries while performing another task. Finally, the agent typically depends on some queries being useful for future interactions, but not necessarily the task at hand. Thus, queries have a higher cost than in traditional active learning setups where the goal of the system is simply to learn a good model.

We examine the usefulness of opportunistic active learning to improve an agent's understanding of natural language descriptions of everyday objects. We consider a task where a robot must identify which member of a set of objects a human user is referring to using natural language. The robot learns classifiers based on multi-sensory information for language predicates that are used to ground natural language descriptions. When trying to understand an object description from one user, the agent is allowed to query for predicate/object labels not directly related to the current interaction.

We compare two agents controlling the robot: one *task-oriented* agent that only asks questions relevant to the current dialog; and one *inquisitive* agent that is willing to ask questions unrelated to the current dialog for expected future performance gains. We show that, in the long run, the inquisitive agent both quantitatively outperforms the task-oriented one at predicting the correct object described by human participants, and is qualitatively rated more fun and usable by participants. To our knowledge, ours is the first work to evaluate the effects of asking off-topic questions to human users interacting with a physical robot performing object identification to improve downstream natural language grounding performance, and, consequently, downstream task performance.[2]

2 Related Work

Most research in active learning is concerned with the design of appropriate metrics to evaluate possible queries' likelihood of improving the current model. Examples include uncertainty sampling [2], density-weighted methods [3], and the presence of conflicting evidence [4]. A survey can be found in [1]. These typically assume that the learner can query any example from the pool of unlabeled examples at any time. In contrast, in our work, the system is restricted so that it can only query for data about a subset of examples (objects) at any time.

Past work compares how human teachers perceive different types of queries a robot may pose during learning from a demonstration task [5]. One notable difference in our work is that the robot in [5] required a human operator to aid in the robot's perception, whereas the system presented in this paper operates autonomously.

Turn-taking interactions where humans have to teach the robot concepts using positive and negative labeled examples are typical for interactive language grounding, but do not employ active learning [6, 7, 8, 9]. Other research uses human-robot interaction, employing forms of active learning to better ground predicates. In those works, the effectiveness of the active learning strategies is not explicitly tested [10, 11, 12, 10], ontological knowledge (pre-coding "red" as a color) is used during grounding [13], or the predicates to be grounded are not drawn from an open-vocabulary of unrestricted user speech [14, 15].

We fill this gap, testing a strategy that asks human users "inquisitive" questions that are off-topic to the task at hand, studying their effects on downstream task performance and human users' perceptions. To our knowledge, ours is the first work to evaluate the effects of asking off-topic questions to human users interacting with physical robots in order to improve natural language grounding. We situate this evaluation within the real-world task of object identification, making the multi-modal, perceptual grounding component a prerequisite, but not ultimate goal for good performance.

3 Object Identification Task and Methods

To test the effectiveness of using active learning to obtain labels for predicates not relevant to the current dialog for long-term rewards, we created an object identification task using a real robot. Figure 1 shows the physical setup of our task. The human participant and robot both started facing Table 2. This table held objects in the active test set O_{te}. The tables flanking the robot (Tables 1 and 3) contained objects in the active training set O_{tr}.

[2]The source code for this experiment is available here: https://github.com/thomason-jesse/perception_classifiers/tree/active_learning

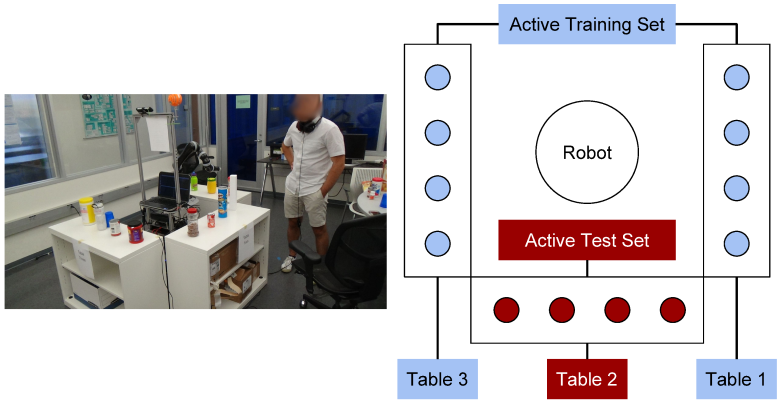


Figure 1: Participants described an object on Table 2 from the *active test set* to the robot in natural language, then answered the robot's questions about the objects in its *active training set* on the side Tables 1 and 3 before the robot guessed the described target object.

Human participants engaged in a dialog with the robot.[3] The robot asked the human to describe one of the four objects in its active test set with a noun phrase. Participants were primed to describe objects with properties, rather than categories, given the motivating example "a fuzzy black rectangle" for "an eraser." Participants were told that the robot had both looked at and interacted with the objects physically using its arm.

Natural Language Grounding. To connect the noun phrases offered by participants to sensory perception, the robot stripped stopwords from the phrase and considered all remaining words as perceptual predicates. We did not restrict the choice of words that participants were allowed to use to describe objects, so our system learned from an open vocabulary. However, it was not equipped to handle multi-word predicates or those that required understanding phrases. The robot then created classifiers to identify these predicates, using objects as positive and negative examples, and getting predicate labels for objects by asking questions about objects in its active training set. Predicates were treated as independent and a separate classifier was learned for each predicate.

Users offered words like "blue," "cylinder," and "heavy" when describing objects. We used a corpus of both visual and non-visual feature representations of objects and their features gathered by multiple interaction behaviors forll2s previous work on an object ordering task [16]. Past work has shown that using non-visual modalities when performing language grounding can help with non-visual words like "heavy," and we follow that work's methodology for training and ensembling SVM classifiers for each predicate to predict whether that predicate applies to a novel object [11]. For every predicate $p \in P$ for P the set of predicates known to the agent and object $o \in O_{tr} \cup O_{te}$, a decision $d(p, o) \in \{-1, 1\}$ and a confidence in that decision are calculated using Cohen's kappa $\kappa(p, o)$ estimated from cross-validation performance on available examples. For predicates with too few examples to train SVMs (at least 2 positive and 2 negative examples were needed to fit the SVMs and obtain confidences), we set $d(p, o) = -1$ and $\kappa(p, o) = 0$ for all objects.

Active Learning Dialog Policy. After the participant described a chosen target object in natural language, the robot asked m questions about objects in its active training set before guessing a target object. Figure 2 gives an overview of this dialog.

With probability q_{yn}, the robot could point to an object and ask the yes/no question "Would you use the word p to describe this object?" for some predicate p. To select the predicate p and object $o \in O_{tr}$ to ask about, we first find the objects in O_{tr} with the lowest confidence κ per predicate (ties broken randomly).

$$o_{\text{min}}(p) = \text{argmin}_{o \in O_{tr}}(\kappa(p, o)).$$

[3] View a demonstration video of the robot system and dialog agents here: https://youtu.be/f-CnIF92_wo

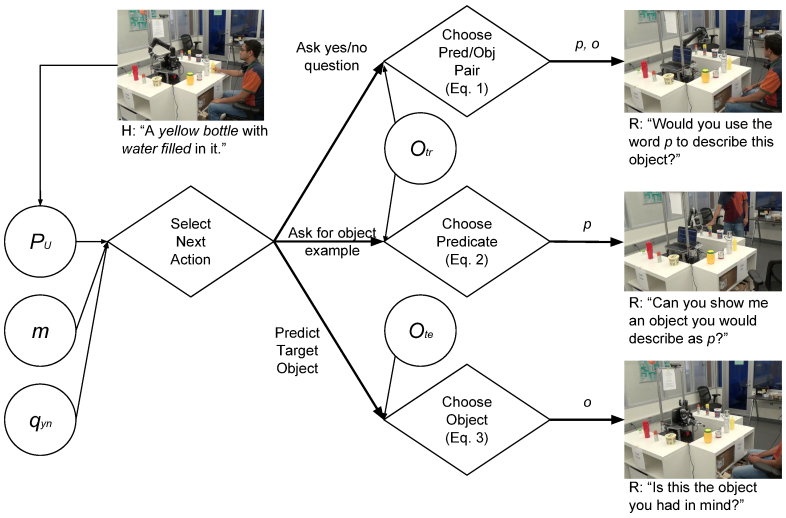


Figure 2: After extracting predicates P_U from a human description of a target object, the dialog agent could ask m questions of two types, choosing one question type over the other with probability q_{yn}. When asking questions, new object labels for the chosen predicate p were restricted to objects in the active training set, $o \in O_{tr}$. After asking m questions, the agent guessed the target object in the active test set $o \in O_{te}$ given its updated grounding classifiers for predicates P_U.

$o_{\text{min}}(p)$ would be the next object whose label is queried in a pool-based active learning setting for p's perceptual classifier with uncertainty sampling as the query strategy [1]. However, as the system is attempting to learn multiple classifiers, it must also choose which of them is to be updated. In order to weight predicates inversely proportional to their confidence in their least confident labels, the predicate p and its corresponding least-confidence object $o_{\text{min}}(p)$, are chosen with probability

$$prob(p) = \frac{1 - \kappa(p, o_{\text{min}}(p))}{\sum_{q \in P \setminus \{p\}} 1 - \kappa(q, o_{\text{min}}(q))}. \tag{1}$$

When querying a predicate and object in the above manner, the robot physically turned to the table holding that object, pointed to it, and asked whether p applied. After getting this new positive or negative example, the robot updated p's perceptual classifier.

With probability $1 - q_{yn}$, the robot instead selected a predicate p and asked "Could you show me an object that you would describe as p?" for which the participant indicated a positive example object for p in the active training set or replied "none," letting the robot know that all objects in the active training set were negative examples of p. A predicate p was selected uniformly at random from those with insufficient data to fit a classifier,

$$p \in \{q : q \in P \wedge \kappa(q, o_{\text{min}}(q)) = 0\}, \tag{2}$$

with the additional constraint that predicates previously asked about in the current dialog with the participant were blacklisted from re-selection. To select an object in the active training set, the participant had to direct the robot to physically face the table containing that object. After getting this new positive label (or, in the case of no object exhibiting p, $|O_{tr}|$ negative labels), the robot updated p's perceptual classifier.

After asking questions, the robot evaluated the predicates $P_U \subseteq P$ of the participant's utterance against the objects in the active test set O_{te}, then turned to the table of candidate objects and pointed to the object that best fit its understanding of the description, $o^* \in O_{te}$,

$$o^* = \text{argmax}_{o \in O_{te}} \left(\sum_{p \in P_U} d(p, o) \kappa(p, o) \right). \tag{3}$$

If the robot guessed incorrectly, the human pointed out the correct object. The target object was considered a positive example for predicates P_U used to describe the object when the agent was fully retrained between rounds (Section 4).

Figure 3: The objects used in our experiments, from fold 0 on the far left to fold 3 on the far right.

Robot Implementation. Experiments were conducted on the BWIBots service robot platform [17]. The robot used a Kinova Mico arm mounted on top of a custom-built mobile base using a Stanley Robotics Segway RMP which rotated to face the three tables holding the objects. The robot used an Asus Xtion Pro RGBD camera mounted at the top of its frame to detect the locations of the objects after turning to face a table, and to detect when a human touched an object. The robot could also reach out and point to an object when asking whether a predicate applied. We implemented robot behaviors in the Robot Operating System,[4] performed text-to-speech using the Festival Speech Synthesis System,[5] and performed automatic-speech-recognition using Google Speech API,[6] recording user speech through a Turtle Beach Ear Force P11 Amplified Stereo Gaming Headset. Once the dialog began, the robot operated autonomously, asking for operator intervention only when the Xtion camera failed to detect the four expected objects on a tabletop surface (often due to being slightly non-orthogonal to a table after a few in-place rotations).

4 Experimental Methodology

We randomly divided the set of 32 objects explored in [16] into 4 folds of 8 objects each, shown in Figure 3. The folds were indexed $\{0, 1, 2, 3\}$.

Two dialog agents controlling the robot were compared. We set policy hyper-parameter $q_{yn} = 0.2$ for both agents. The *baseline* agent was only allowed to ask questions about the predicates relevant to the current dialog. That is, if a person described the target object as "a blue cylinder," then the *baseline* agent could only ask about "blue" and "cylinder." We set $m = 3$ for the *baseline* agent. By contrast, the *inquisitive* agent was allowed to ask questions about any predicate it had previously heard. Thus, the *inquisitive* agent could ask about "heavy" even if a user used "a blue cylinder" to describe the target object. We set $m = 5$ for the *inquisitive* agent, making it both more talkative and less task-oriented than the *baseline* agent. It would be interesting to learn the optimal values of the parameters for task performance via reinforcement learning, which we leave for future work. In the final round of testing described below, the *inquisitive* agent was restricted to on-topic questions and had $m = 3$, making the agents differ only by their training strategies up to that point.

In the object identification game, the objects in the active training set were from fold n when the objects in the active test set were in fold $n + 1$. Both agents began knowing no predicates. Each human participant was assigned to one of the two agents before the session began. Participants played two games each. Both games had the 8 objects in the active training set randomly ordered on the robot's side tables. Between games, the 4 objects on the robot's front table were alternated such that all 8 of the objects in the active test set had a chance to be described by the participant. After the games, each participant filled out an exit survey, discussed in Section 5. The agents were tested across three **rounds**, with objects from the active test set moving to the active training set between rounds. The rounds can be summarized by:

- **Round 1** with fold 0 as the active training set and fold 1 the active test set, the agents effectively differed only by the number of questions they could ask, since neither had seen any predicates before;

[4] http://www.ros.org/

[5] http://www.cstr.ed.ac.uk/projects/festival/

[6] https://cloud.google.com/speech/

- **Round 2** with fold 1 as the active training set and fold 2 the active test set, the *inquisitive* agent could ask about the predicates in the current dialog or any predicate it learned from round 1;
- **Round 3** with fold 2 as the active training set and fold 3 the active test set, the agents both operated by the *baseline* rules (on-topic questions, $m = 3$), comparing the effects of the training strategies used in rounds 1 and 2.

Between rounds, the dialogs that the agents had with their participants were aggregated and new predicate classifiers were trained to use in the next round. The agents were trained independently, with the *baseline* agent only using conversations it had, and the *inquisitive* agent only using conversations it had. This training aggregation was done in round-based batches so that the objects in the *active test set* were always unseen by the agents' trained classifiers at the time of any given conversation.

We hypothesized that:

1. The *inquisitive* agent would guess the correct object more often than the *baseline* agent.
2. Users would not qualitatively dislike the *inquisitive* agent for asking too many questions and being off-topic compared to the *baseline* agent.

5 Experimental Results

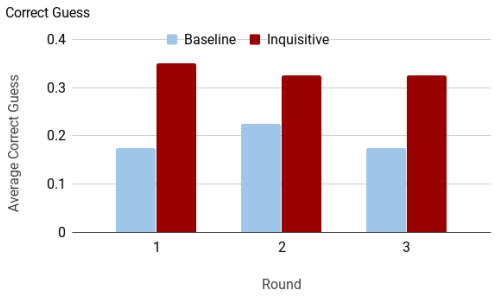
Five participants played two games each with the robot for each agent in each round. In total, 30 study participants, comprising graduate and undergraduate students and employees at our university, interacted with the robot. After two games, participants filled out an exit survey by answering "Strongly Disagree," "Disagree," "Neutral," "Agree," or "Strongly Agree" (mapped to scores $0 - 4$) to the questions shown in Figure 4.

Results. Figure 4 shows the robot's average correctness across rounds between the two agents, as well as the average results of exit survey questions. The *inquisitive* agent consistently outperforms the *baseline* agent at identifying the correct object (Figure 4a).[7] In round 3, when the agents were both restricted to $m = 3$ questions and only on-topic predicates, the difference in performance is entirely attributable to the training strategies of the agents so far, and the *inquisitive* agent again has a higher rate of predicting the target object. The *inquisitive* agent outperforms a random chance baseline (0.25 average correctness for 4 objects), while the *baseline* agent performs slightly worse due to noisy perceptual classifiers with few positive and negative examples. The *inquisitive* agent is perceived as more understanding on average than the *baseline* agent (Figure 4b). These results support our hypothesis that the *inquisitive* agent would outperform the *baseline* agent at the object identification task.

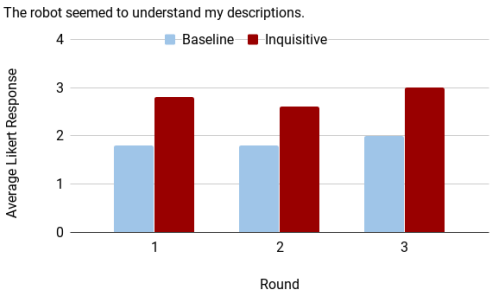
The *inquisitive* agent is perceived as asking too many questions slightly more often than the *baseline* agent in round 2, when it can ask about predicates not related to the current dialog, but not in round 1, where it still asks 2 more questions than the *baseline* agent on average, but they are on-topic (Figure 4c). The trends are nearly identical for the similar question of whether users felt the conversation went on too long (Figure 4f). The *inquisitive* agent scored higher with human users across rounds than the *baseline* agent for the prompts about whether the robot was fun (Figure 4d) and whether the user would use a household robot like this one to get objects in another room (Figure 4e). These results support our hypotheses that the *inquisitive* agent would not be disliked for asking too many questions or being off-topic.

Discussion. The *inquisitive* agent's differences from the *baseline* agent in round 3 partially rely on predicates from previous rounds being used again in that round. In general, asking about an off-topic predicate only helps if that predicate will be seen again in the future. Table 1 shows the predicates introduced in each round as well as those repeated from a previous round. There is substantial overlap, indicating that the dataset of objects used is homogeneous enough that learning predicates from previous folds is helpful when identifying objects in unseen folds. Additionally,

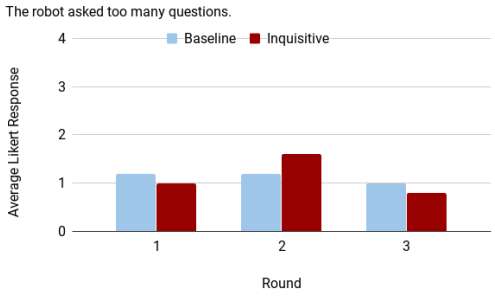
[7] We note that in the event of T tied confidences for an object to select, with the correct object among those tied, we reward the robot $\frac{1}{|T|}$ correctness, regardless of the random choice it made among those T.



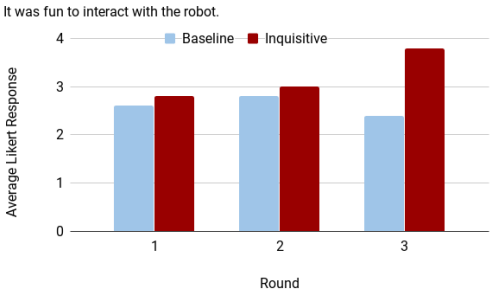
(a) Correct guesses



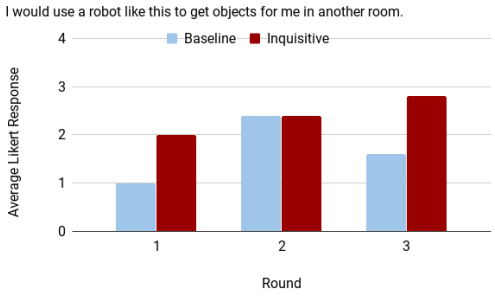
(b) "The robot seemed to understand my descriptions."



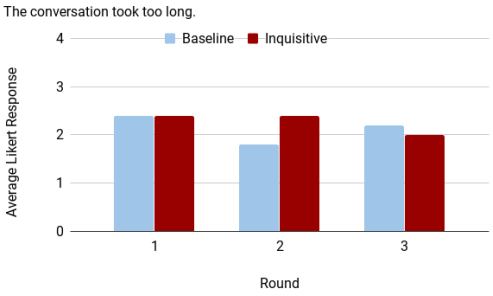
(c) "The robot asked too many questions."



(d) "It was fun to interact with the robot."



(e) "I would use a robot like this to get objects for me in another room."



(f) "The conversation took too long."

Figure 4: Comparing average robot correct guess and average user survey responses across the three rounds between the two agents. In round 1, the agents differed only in the number of questions they could ask. In round 2, the *inquisitive* agent could both ask more questions and ask about off-topic predicates. In round 3, the agents differed only in their training so far, and both had number of questions fixed to $m = 3$.

First seen	Predicates Used by Round		
	round 1	round 2	round 3
round 1	71	32	32
round 2		37	14
round 3			24
Total	**71**	**69**	**70**

Table 1: The number of unique predicates introduced in each round and repeated in subsequent rounds. The diagonal shows predicates used for the first time in each round, while the bottom row shows the total unique predicates used (regardless of when they were first seen) per round.

there were 132 unique predicates introduced throughout all thirty participants' games, suggesting that the dataset is diverse enough to elicit a wide range of language predicates. Descriptions varied in length among users, as well, from 1 predicate in utterances like "*heavy*" to 8 predicates in the utterance "a *transparent plastic bottle* with *brown peanuts inside* it with the *red cap*."

6 Conclusions and Future Work

We introduce opportunistic active learning, where a system engaged in a task makes use of active learning metrics to query for labels potentially useful for future tasks. We demonstrate that a robot using opportunistic active learning during an object identification task performs well in understanding unrestricted natural language descriptions of a target object. Our robot experiments simulate a household robot that can be used to retrieve distant objects and is allowed to first ask questions about nearby objects to help clarify its understanding of natural language predicates.

The robot can ask humans not just questions about words relevant for the current task (e.g. questions about "blue" and "cylinder" when told to "go get me the blue cylinder") but about any words it currently understands poorly. We demonstrate that such an *inquisitive* agent not only outperforms an agent that stays on-topic with its questions at identifying the correct object described by a human user, but that users find the *inquisitive* agent, on average, more comprehending, fun, and usable in a real-world setting.

We are interested in learning a dialog policy that allows inquisitive active learning, but balances it against current task success, rather than explicitly limiting the robot to a certain number of questions per dialog as in this work. Additionally, with more objects, more rounds of training could be done to explore whether there is a ceiling on performance gains from asking off-topic questions. There has been some work on task-oriented dialog learning for robotics, but that work did not incorporate language grounding with perceptual predicates [18], did not involve much initiative from the robot [19], or aimed at eliciting human collaboration required only for completing the current task [20].

Finally, by performing more complicated language understanding than simple stopword removal, future work may also be able to learn to jointly understand language predicates and human commands (as in [21, 22, 23]), additionally leveraging opportunistic active learning to ask questions that best clarify both perceptual grounding and semantic understanding.

Acknowledgments

This work is supported by a National Science Foundation Graduate Research Fellowship to the first author, an NSF EAGER grant (IIS-1548567), and an NSF NRI grant (IIS-1637736). A portion of this work has taken place in the Learning Agents Research Group (LARG) at UT Austin. LARG research is supported in part by NSF (CNS-1305287, IIS-1637736, IIS-1651089, IIS-1724157), TxDOT, Intel, Raytheon, and Lockheed Martin. Peter Stone serves on the Board of Directors of Cogitai, Inc. The terms of this arrangement have been reviewed and approved by the University of Texas at Austin in accordance with its policy on objectivity in research.

References

[1] B. Settles. Active learning literature survey. *University of Wisconsin, Madison*, 52(55-66):11, 2010.

[2] D. D. Lewis and W. A. Gale. A sequential algorithm for training text classifiers. In *Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 3–12. Springer-Verlag New York, Inc., 1994.

[3] B. Settles and M. Craven. An analysis of active learning strategies for sequence labeling tasks. In *Proceedings of the conference on empirical methods in natural language processing*, pages 1070–1079. Association for Computational Linguistics, 2008.

[4] M. Sharma and M. Bilgic. Evidence-based uncertainty sampling for active learning. *Data Mining and Knowledge Discovery*, 31(1):164–202, 2017.

[5] M. Cakmak and A. L. Thomaz. Designing robot learners that ask good questions. In *Human-Robot Interaction (HRI), 2012 7th ACM/IEEE International Conference on*, pages 17–24, March 2012.

[6] H. Dindo and D. Zambuto. A probabilistic approach to learning a visually grounded language model through human-robot interaction. In *International Conference on Intelligent Robots and Systems*, pages 760–796, Taipei, Taiwan, 2010. IEEE.

[7] T. Kollar, J. Krishnamurthy, and G. Strimel. Toward interactive grounded language acquisition. In *Robotics: Science and Systems*, 2013.

[8] N. Parde, A. Hair, M. Papakostas, K. Tsiakas, M. Dagioglou, V. Karkaletsis, and R. D. Nielsen. Grounding the meaning of words through vision and interactive gameplay. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence*, pages 1895–1901, Buenos Aires, Argentina, 2015.

[9] I. Lutkebohle, J. Peltason, L. Schillingmann, B. Wrede, S. Wachsmuth, C. Elbrechter, and R. Haschke. The curious robot-structuring interactive robot learning. In *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*, pages 4156–4162. IEEE, 2009.

[10] A. Vogel, K. Raghunathan, and D. Jurafsky. Eye spy: Improving vision through dialog. In *Association for the Advancement of Artificial Intelligence*, pages 175–176, 2010.

[11] J. Thomason, J. Sinapov, M. Svetlik, P. Stone, and R. Mooney. Learning multi-modal grounded linguistic semantics by playing "I spy". In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 3477–3483, July 2016.

[12] D. Skočaj, A. Vrečko, M. Mahnič, M. Janíček, G.-J. M. Kruijff, M. Hanheide, N. Hawes, J. L. Wyatt, T. Keller, K. Zhou, et al. An integrated system for interactive continuous learning of categorical knowledge. *Journal of Experimental & Theoretical Artificial Intelligence*, 28(5): 823–848, 2016.

[13] S. Mohan, A. H. Mininger, J. R. Kirk, and J. E. Laird. Acquiring grounded representations of words with situated interactive instruction. In *Advances in Cognitive Systems*, 2012.

[14] M. Cakmak, C. Chao, and A. L. Thomaz. Designing interactions for robot active learners. *IEEE Transactions on Autonomous Mental Development*, 2(2):108–118, June 2010. ISSN 1943-0604.

[15] J. Kulick, M. Toussaint, T. Lang, and M. Lopes. Active learning for teaching a robot grounded relational symbols. In *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, pages 1451–1457. AAAI Press, 2013. ISBN 978-1-57735-633-2.

[16] J. Sinapov, P. Khante, M. Svetlik, and P. Stone. Learning to order objects using haptic and proprioceptive exploratory behaviors. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence*, 2016.

[17] P. Khandelwal, S. Zhang, J. Sinapov, M. Leonetti, J. Thomason, F. Yang, I. Gori, M. Svetlik, P. Khante, V. Lifschitz, J. K. Aggarwal, R. Mooney, and P. Stone. BWIbots: A platform for bridging the gap between AI and human–robot interaction research. *The International Journal of Robotics Research*, 2017. doi:10.1177/0278364916688949.

[18] A. Padmakumar, J. Thomason, and R. J. Mooney. Integrated learning of dialog strategies and semantic parsing. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pages 547–557, Valencia, Spain, April 2017.

[19] L. She, S. Yang, Y. Cheng, Y. Jia, J. Chai, and N. Xi. Back to the Blocks World: Learning New Actions through Situated Human-Robot Dialogue. In *Proceedings of the 15th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)*, pages 89–97, 2014.

[20] S. Tellex, R. A. Knepper, A. Li, N. Roy, and D. Rus. Asking for Help Using Inverse Semantics. In *Proceedings of the 2016 Robotics: Science and Systems Conference (RSS)*, 2014.

[21] T. Kollar, V. Perera, D. Nardi, and M. Veloso. Learning Environmental Knowledge from Task-Based Human-Robot Dialog. In *Proceedings of the 2013 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4304–4309, 2013.

[22] Y. Bisk, D. Yuret, and D. Marcu. Natural language communication with robots. In *Proceedings of the 15th Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 751–761, San Diego, CA, June 2016.

[23] M. Alomari, P. Duckworth, D. C. Hogg, and A. G. Cohn. Natural language acquisition and grounding for embodied robotic systems. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, pages 4349–4356, February 2017.