Interactive Learning of State Representation through Natural Language Instruction and Explanation

Qiaozi Gao, Lanbo She, and Joyce Y. Chai
Department of Computer Science and Engineering
Michigan State University, East Lansing, MI 48824, USA
{gaoqiaoz, shelanbo, jchai}@cse.msu.edu

Abstract

One significant simplification in most previous work on robot learning is the closed-world assumption where the robot is assumed to know ahead of time a complete set of predicates describing the state of the physical world. However, robots are not likely to have a complete model of the world especially when learning a new task. To address this problem, this extended abstract gives a brief introduction to our on-going work that aims to enable the robot to acquire new state representations through language communication with humans.

Introduction

As cognitive robots start to enter our lives, being able to teach robots new tasks through natural interaction becomes important (Matuszek et al. 2012; Liu et al. 2016a; 2016b; Chai, Cakmak, and Sidner 2017). One of the most natural ways for humans to teach task knowledge is through natural language instructions, which are often expressed by verbs or verb phrases. Previous work has investigated how to connect action verbs to low-level primitive actions (Branavan et al. 2009; Mohan and Laird 2014; She et al. 2014; Misra et al. 2015; 2016; She and Chai 2016; 2017). In most of these studies, a robot first acquires the state change of an action from human demonstrations and represents verb semantics using the desired goal state. With learned verb semantics, given a language instruction, the robot can apply the goal states of the involved verbs to plan for a sequence of low-level actions.

For example, a human can teach the robot the meaning of the verb phrase *"heat water"* through step-by-step instructions as shown in H2 in Figure 1. The robot can identify the state change by comparing the final environment to the initial environment. The learned verb semantics is represented by the goal state (e.g., `Temp(x,High)`). To handle uncertainties of perception, the robot can also ask questions and acquire better representations of the world through interaction with humans (She and Chai 2017).

Previous work is developed based on a significant simplification: the robot knows ahead of time a complete set of predicates (or classifiers) that can describe the state of the physical world. However in reality robots are not likely
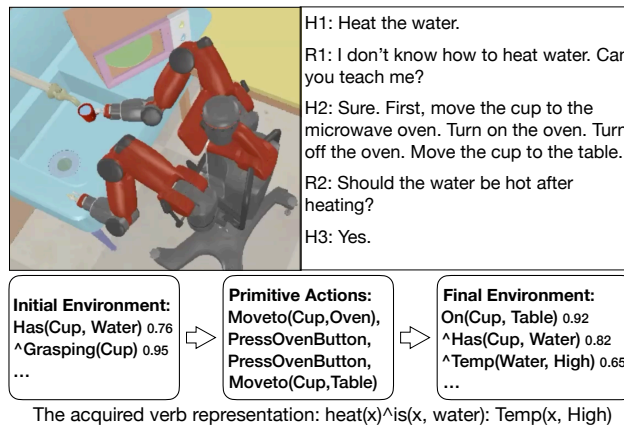
H1: Heat the water.

R1: I don't know how to heat water. Can you teach me?

H2: Sure. First, move the cup to the microwave oven. Turn on the oven. Turn off the oven. Move the cup to the table.

R2: Should the water be hot after heating?

H3: Yes.

Initial Environment:
Has(Cup, Water) 0.76
^Grasping(Cup) 0.95
...

Primitive Actions:
Moveto(Cup,Oven),
PressOvenButton,
PressOvenButton,
Moveto(Cup,Table)

Final Environment:
On(Cup, Table) 0.92
^Has(Cup, Water) 0.82
^Temp(Water, High) 0.65
...

The acquired verb representation: heat(x)^is(x, water): Temp(x, High)

Figure 1: An example of learning the state-based representation for the command *"heat water"*.

to have a complete model of the world. Thus, it is important for the robot to be proactive (Chai et al. 2014; 2016) and transparent (Alexandrova et al. 2014; Alexandrova, Tatlock, and Cakmak 2015; Whitney et al. 2016; Hayes and Shah 2017) about its internal representations so that humans can provide the right kind of feedback to help capture new world states. To address this problem, we are developing a framework that allows the robot to acquire new states through language communication with humans.

Interactive State Acquisition

The proposed framework is shown in Figure 2. In additional to modules to support language communication (e.g., **grounded language understanding** and **dialogue manager**) and action (e.g., **action planning** and **action execution**), the robot has a **knowledge base** and a **memory/experience** component. The **knowledge base** contains the robot's existing knowledge about verb semantics, state predicates, and action schema (both primitive actions and high-level actions). The **memory/experience** component keeps track of interaction history such as language input from the human and sensory input from the environment.

Suppose the robot does not have the state predicate `Temp(x, High)` in its knowledge base and the effect of the primitive action `PressOvenButton` only describes
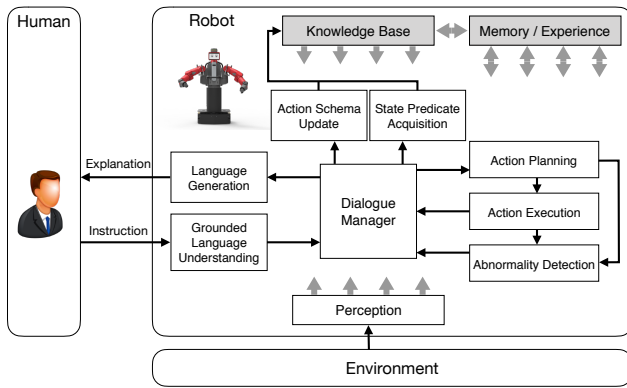
Figure 2: Interactive acquisition of new physical states.

H1: Now I will teach you how to heat water. First, move the cup to the oven. Turn on the oven. Turn off the oven. Move the cup to the table.
R1: According to my knowledge, I can achieve the same goal without moving the cup to the oven and pressing the oven button. What is the effect of these extra actions?
H2: By doing these actions, the temperature of water will be high.
R2: Is this effect caused by pressing the oven button?
H3: Yes.
R3: Ok. [Add a new predicate *Temp(x, High)* to the knowledge base.]
...
R4: In order to increase the temperature of water, is it necessary to put water in the oven before pressing the oven button?
H4: Yes, it is necessary.
R5: OK. [Add a pair of condition and effect to the primitive action *PressOvenButton*: *if In(x, Oven), then Temp(x, High)*]
...

Figure 3: An example of interactively learning a new state predicate during the human teaches the robot how to "heat water".

the change of the oven status (i.e., `Status(Oven, On)`). Our framework will allow the robot to acquire the new state predicate `Temp(x, High)` and update action representation (shown below with the added condition and state in bold) through interaction with the human as shown in Figure 3.

if (not Status(Oven, On)), then:
    Status(Oven, On) **and if In(x, Oven), then: Temp(x, High)**
if Status(Oven, On), then:
    not Status(Oven, On)

This framework includes two main processes: (1) acquiring and detecting new states; and (2) updating action representation.

Acquiring and Detecting New States

Since an incomplete action schema can cause planning problems (Gil 1994), the robot can potentially discover the related abnormality by retrospective planning. In our example, the robot does not have the state predicate `Temp(x,High)` in its current knowledge base. Thus in the robot's mind, the final environment will not contain `Temp(Water, High)`. After the human provides instructions on how to heat water, the dialogue manager calls a retrospective planning process based on the robot's current knowledge to achieve the final environment. Then the **abnormality detection** module compares the planned action sequence with human provided action sequence and finds that the planning result lacks of primitive actions `Moveto(Cup, Oven)` and `PressOvenButton`. Once an abnormality is detected, the robot explains its limitation to human for diagnosis (R1). Note that there is a gap between the robot's mind and the human's mind. The human does not know the state predicates that the robot uses to represent the physical world. In order for humans to understand its limitation, the robot explains the differences between the two action sequences, and requests the human to provide missing effects. Based on the human's response, the **state predicate acquisition** module adds a new state predicate `Temp(x, High)` to the knowledge base. Next the robot needs to know how to detect such state from the physical environment. State detection is a challenging problem by itself. It often involves classifying continuous signals from the sen-

sors into certain classes, for examples, as in previous work that jointly learns concepts and their physical groundings by integrating language and vision (Matuszek et al. 2012; Krishnamurthy and Kollar 2013). We are currently exploring approaches that automatically bootstrap training examples from the web for detection of state.

Updating Action Representation

Once a new state predicate is acquired, the robot needs to know what primitive actions and under what conditions the related state change can be caused. The relevant primitive action can be identified by applying the state detection model to the sensory input from the environment that is stored in the memory. Now the problem is reduced to determine what condition is needed to cause that particular state change. And this is similar to the *planning operator acquisition* problem, which has been studied extensively (Wang 1995; Amir and Chang 2008; Mourão et al. 2012; Zhuo and Yang 2014). However, in previous work, primitive actions are acquired based on multiple demonstration instances. Inspired by recent work that support interactive question answering (Cakmak and Thomaz 2012; She and Chai 2017), we intend to enable robots to ask questions to identify the correct conditions for primitive actions (R4). We are currently extending an approach based on reinforcement learning to learn when to ask what questions. Based on the human's response, the **action schema update** module adds a pair of condition and effect to the primitive action `PressOvenButton` as shown earlier.

Conclusion and Future Work

This paper gives a brief introduction to our on-going work that enables the robot to acquire new state predicates to better represent the physical world through language communication with humans. Our current and future work is to evaluate this framework in both offline data and real-time interactions, and extend it to interactive task learning.

Acknowledgment

References

Alexandrova, S.; Cakmak, M.; Hsaio, K.; and Takayama, L. 2014. Robot programming by demonstration with interactive action visualizations. In *In Robotics: Science and Systems (RSS)*.

Alexandrova, S.; Tatlock, Z.; and Cakmak, M. 2015. Roboflow: A flow-based visual programming language for mobile manipulation tasks. In *IEEE International Conference on Robotics and Automation (ICRA)*.

Amir, E., and Chang, A. 2008. Learning partially observable deterministic action models. *Journal of Artificial Intelligence Research* 33:349–402.

Branavan, S.; Chen, H.; Zettlemoyer, L. S.; and Barzilay, R. 2009. Reinforcement learning for mapping instructions to actions. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 1-Volume 1*, 82–90. Association for Computational Linguistics.

Cakmak, M., and Thomaz, A. L. 2012. Designing robot learners that ask good questions. In *ACM/IEEE International Conference on Human-Robot Interaction*, 17–24.

Chai, J. Y.; She, L.; Fang, R.; Ottarson, S.; Littley, C.; Liu, C.; and Hanson, K. 2014. Collaborative effort towards common ground in situated human-robot dialogue. In *The 9th ACM/IEEE Conference on Human-Robot Interaction (HRI)*.

Chai, J. Y.; Fang, R.; Liu, C.; and She, L. 2016. Collaborative language grounding towards situated human robot dialogue. *AI Magazine* 37(4):32–45.

Chai, J. Y.; Cakmak, M.; and Sidner, C. 2017. Teaching robots new tasks through natural interaction. In Gluck, K. A., and Laird, J. E., eds., *Interactive Task Learning: Agents, Robots, and Humans Acquiring New Tasks through Natural Interactions, Strüngmann Forum Reports, J. Lupp, series editor*, volume 26. MIT Press.

Gil, Y. 1994. Learning by experimentation: Incremental refinement of incomplete planning domains. In *International Conference on Machine Learning*, 87–95.

Hayes, B., and Shah, J. 2017. Improving robot controller interpretability and transparency through autonomous policy explanation. In *ACM International Conference on Human-Robot Interaction*.

Krishnamurthy, J., and Kollar, T. 2013. Jointly learning to parse and perceive: Connecting natural language to the physical world. *Transactions of the Association for Computational Linguistics* 1:193–206.

Liu, C.; Chai, J. Y.; Shukla, N.; and Zhu, S. 2016a. Task learning through visual demonstration and situated dialogue. In *AAAI 2016 Workshop on Symbiotic Cognitive Systems*.

Liu, C.; Yang, S.; Saba-Sadiya, S.; Shukla, N.; He, Y.; Zhu, S.-C.; and Chai, J. 2016b. Jointly learning grounded task structures from language instruction and visual demonstration. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 1482–1492.

Matuszek, C.; Fitzgerald, N.; Zettlemoyer, L.; Bo, L.; and Fox, D. 2012. A joint model of language and perception for grounded attribute learning. In *Proceedings of the 29th International Conference on Machine Learning (ICML-12)*, 1671–1678.

Misra, D. K.; Tao, K.; Liang, P.; and Saxena, A. 2015. Environment-driven lexicon induction for high-level instructions. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics*, 992–1002.

Misra, D. K.; Sung, J.; Lee, K.; and Saxena, A. 2016. Tell me dave: Context-sensitive grounding of natural language to manipulation instructions. *The International Journal of Robotics Research* 35(1-3):281–300.

Mohan, S., and Laird, J. E. 2014. Learning goal-oriented hierarchical tasks from situated interactive instruction. In *AAAI*, 387–394.

Mourão, K.; Zettlemoyer, L. S.; Petrick, R. P. A.; and Steedman, M. 2012. Learning STRIPS operators from noisy and incomplete observations. In *Proceedings of the Twenty-Eighth Conference on Uncertainty in Artificial Intelligence*, 614–623.

She, L., and Chai, J. Y. 2016. Incremental acquisition of verb hypothesis space towards physical world interaction. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, volume 1.

She, L., and Chai, J. Y. 2017. Interactive learning of grounded verb semantics towards human-robot communication. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, volume 1.

She, L.; Yang, S.; Cheng, Y.; Jia, Y.; Chai, J.; and Xi, N. 2014. Back to the blocks world: Learning new actions through situated human-robot dialogue. In *Proceedings of the SIGDIAL 2014 Conference*.

Wang, X. 1995. Learning by observation and practice: An incremental approach for planning operator acquisition. In *ICML*, 549–557.

Whitney, D.; Eldon, M.; Oberlin, J.; and Tellex, S. 2016. Interpreting multimodal referring expressions in real time. In *IEEE International Conference on Robotics and Automation (ICRA)*, 3331–3338.

Zhuo, H. H., and Yang, Q. 2014. Action-model acquisition for planning via transfer learning. *Artificial intelligence* 212:80–103.