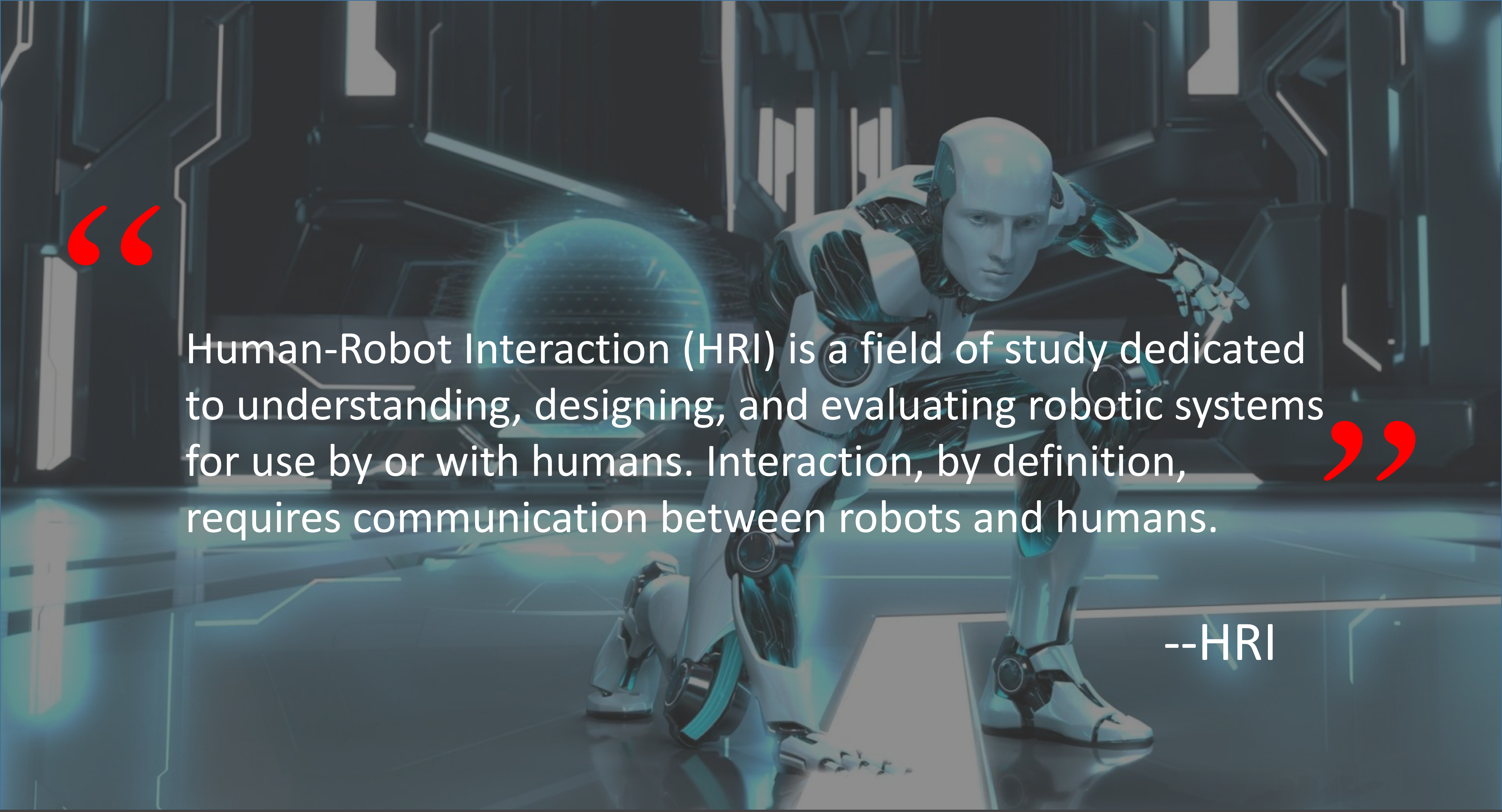
Interactive Learning of Grounded Verb Semantics towards Human-Robot Communication

Lanbo She and Joyce Y. Chai
Department of Computer Science and Engineering
Michigan State University

Presenter: Yuyang Rao
April 2017



“Human-Robot Interaction (HRI) is a field of study dedicated to understanding, designing, and evaluating robotic systems for use by or with humans. Interaction, by definition, requires communication between robots and humans.”

--HRI

Overview

Interactive Learning of Grounded Verb Semantics towards Human-Robot Communication

Verb Semantics

- Command eg: Boil the water

Interactive Learning

- Questions
- Answers

Goal

- Better Human robot interaction

Introduction

Challenge

robots do not have sufficient linguistic or world knowledge as humans do

Interactive learning

allows robots to proactively engage in interaction with human partners

Reinforcement learning

Reward system + update knowledge base

Previous Work

Learning Approach

Learning Reply on on multiple instances of human demonstrations of corresponding actions.

Disadvantage 1

Under the assumption of perfect perception of the environment. However, does not hold in real-world situated interaction.

Disadvantage 2

Each demonstration is simply a sequence of primitive actions associated with a verb. No other type of interaction

State-based Representation

How Robot understand the command: “boil the water”

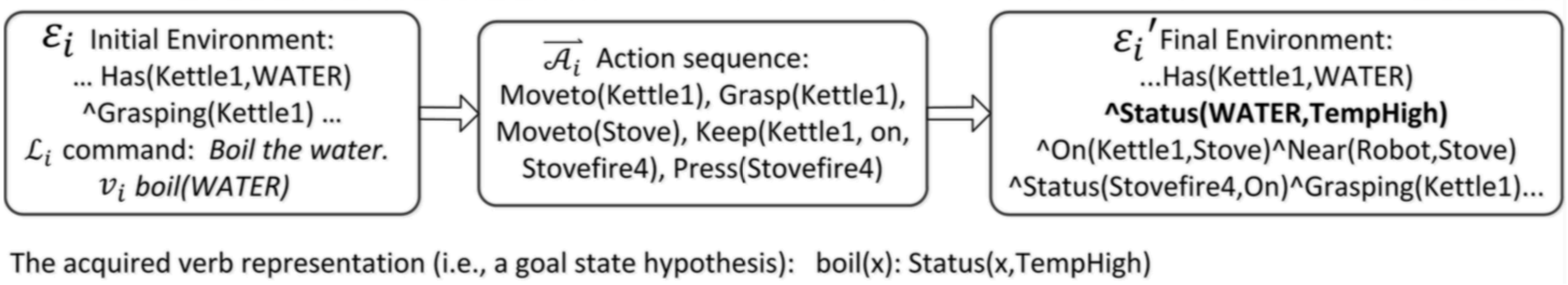


Figure 1: An example of acquiring state-based representation for verb semantics based on an initial environment $\mathcal{E}_i$, and a language command $\mathcal{L}_i$, the primitive action sequence $\overrightarrow{\mathcal{A}_i}$ demonstrated by the human, and the final environment $\mathcal{E}'_i$ that results from the execution of $\overrightarrow{\mathcal{A}_i}$ in $\mathcal{E}_i$.

Hypothesis Space

Command

Boil the water (verb phrase)

Execution

select a most relevant hypothesis and use the corresponding goal state to plan for actions to execute.

Learning

If fails, ask the human for a demonstration.

Update

Based on the demonstrated actions, the robot will learn a new representation

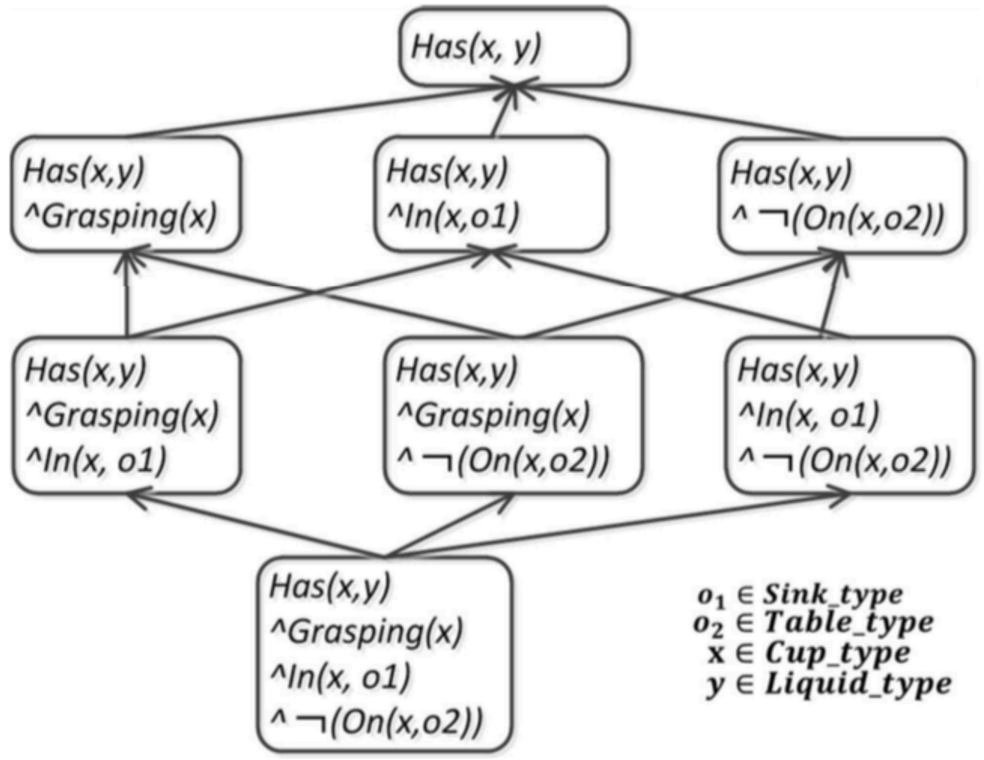


Figure 2: An example hypothesis space for the verb frame $fill(x, y)$.

Probabilistic Environment:

...^Has(Kettle1,Water) 0.64 ^Grasping(Kettle1) 0.91
^Status(Kettle1,HighTemp) 0.95 ^On(Kettle1,Stove) 0.2
^Near(Robot,Stove) 0.43 ^Status(Stovefire4,On) 0.6 ^...

Figure 3: An example probabilistic sensing result.

Environment representation

The environment representation is often partial, error prone, and full of uncertainties.

- [x] Yes
- [ ] No

Interactive learning approach aims to address these uncertainties

Previous works assume: perfect, deterministic representation

Framework of Interactive Learning

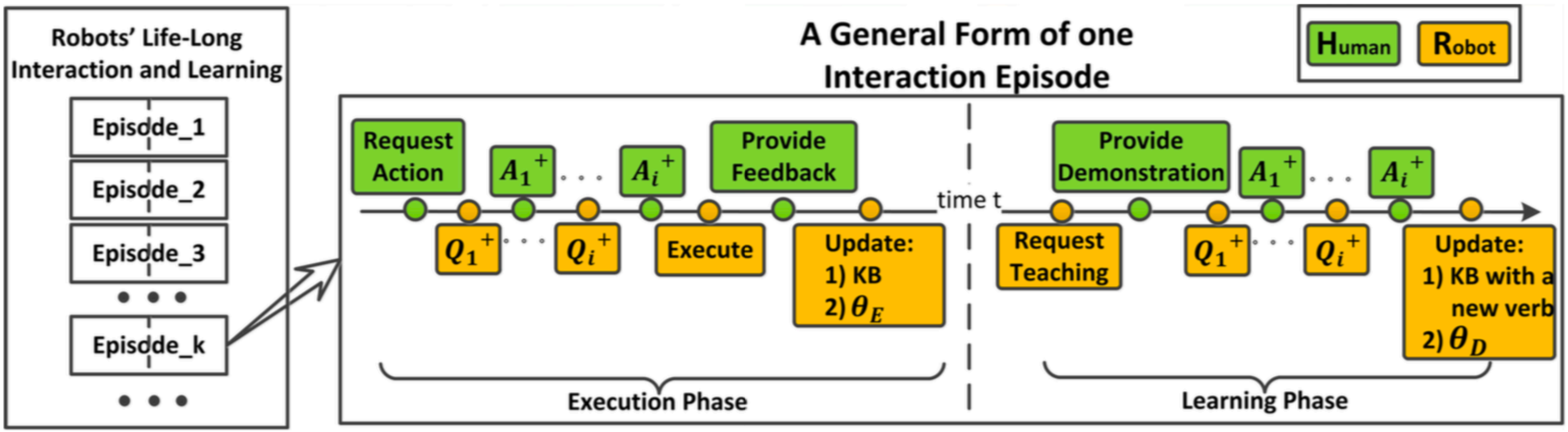


Figure 4: A general framework of robot interactive learning. KB stands for knowledge base, θ_E stands for Interaction Strategy for Execution, and θ_D stands for Interaction Strategy for Learning.

The Big Question

When to

Ask

What

Question ?

What Questions to Ask?

Action Name	Explanation	Question Example	Reward
1. **np_grd_whq**(n)	Ask for the grounding of a np.	*"Which is the cup, can you show me?"*	-6.5[1]
2. **np_grd_ynq**(n, o)	Confirm the grounding of a np.	*"I see a silver object, is that the pot?"*	-1.0 / -2.0
3. **env_pred_ynq**(p)	Confirm a predicate in current environment.	*"Is the microwave door open?"*	-1.0 / -2.0
4. **goal_pred_ynq**(p)	Confirm whether a predicate p should be in the final environment.	*"Is it true the pot should be on the counter?"*	-1.0 / -2.0
5. **select_hypo**(h)	Choose a hypothesis to use as goal and execute.		100 / -2.0
6. **bulk_np_grd_ynq**(n, o)	Confirm the grounding of multiple nps.	*"I think the pot is the red object and milk is in the white box, am I right?"*	-3.0 / -6.0[2]
7. **pred_change_ynq**(p)	Ask whether a predicate p has been changed by the action demonstration.	*"The pot is on a stand after the action, is that correct?"*	-1.0 / -2.0
8. **include_fluent**($\wedge p$)	Include $\wedge p$ into the goal state representation. Update the verb semantic knowledge.		100 / -2.0

When to?

Algorithm

1. **Input**
   e, l, Feature function ϕ;
   Old policy θ ; Hypotheses H;

2. **Initialize**
   state s initialized with e, l ;
   first action a ~ P (a|s; θ)

3. while s *is not terminal* do
   Take action a, receive reward r;
   update State s
   Choose a' ~ P(a'|s';θ)
   End

4. if *s terminates with positive feedback* then
   Update H;
   end

5. Output : Updated H and θ.

Execution Example

H1: Boil the water.
R1: (*select a hypothesis and execute the action accordingly*) H2: (*provide true/false feedback on whether the execution is correct*)

Execution Example

H1: Boil the water.
R1: I see a silver object on the table, is that a pot? H2: Yes.
R2: Should I leave the pot on the stove after boil? H3: Yes.
R3: Ok. (*select a hypothesis to execute*)
H4: Thank you.

Comparison

Learning Example

H3: Boil the water.
R3: I don't know how to boil, could you show me an example? H4: (*human shows a sequence of primitive actions*)
R4: (*indicate the acceptance and induce a hypothesis to represent verb semantics*)

Learning Example

H5: Boil the water.
R5: I don't know how to boil, could you show me an example? H6: Ok. (*demonstrate the task, at the end put the pot on a heat pad*)
R6: So after boiling something, should I put it on a heat pad? H7: No, you don't have to.
R7: I see.

Experiment Setup

Dataset

(1) Kitchen & living room;
(2) 979 instances

Noisy Environment Representation

(1)PerfectEnv
(2)NormStd3
(3)NormStd5
(4)UniEnv

Evaluation Metrics

(1) *IED:* action sequence
(2) *SJI:* state changes

Approaches

(1) She 16
(2) *RandomPolicy*
(3) *ManualPolicy*

Result

1. The interactive learning with RL policy outperforms the previous approach *She16*.
2. The RL policy slightly outperforms interactive learning using manually defined policy.
3. However, *the manualPolicy* results in much longer interaction (i.e., more questions) than the RL policy.

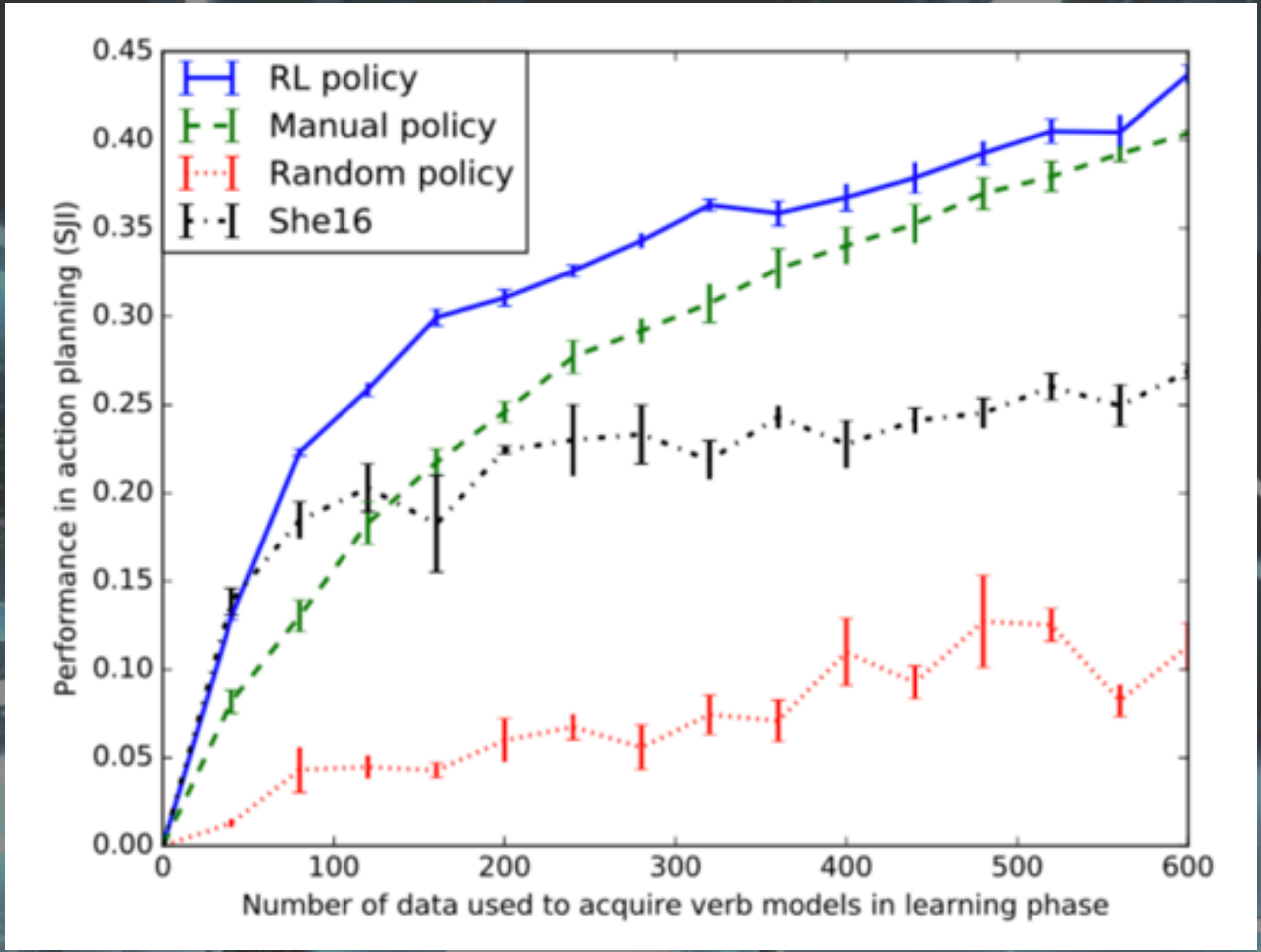


Figure 5: Performance (SJI) comparison on different interaction policies to the testing data.

Result

1. When the environment becomes noisy, the performance of *She16* that only relies on demonstrations decreases significantly.
2. IL improves the performance under the perfect environment condition
3. Effect in noisy environment is more remarkable.

	She16		*RL policy*		% improvement	
	IED	*SJI*	*IED*	*SJI*	*IED*	*SJI*
PerfectEnv	0.430	0.426	0.453	0.468	5.3%*	9.9%*
NormStd3	0.284	0.273	0.420	0.431	47.9%*	57.9%*
NormStd5	0.172	0.168	0.392	0.411	127.9%*	144.6%*
UniEnv	0.168	0.163	0.332	0.347	97.6%*	112.9%*

Table 1: Performance comparison between *She16* and our interactive learning based on environment representations with different levels of noise

Conclusion

Now

Robots live in a noisy environment, full of uncertainties.

Asking intelligent questions to interact with human can handle the uncertainties

Future Work

To learn new predicates by interaction with humans

Deep neural network to alleviate feature engineering


THANK YOU
ANY QUESTIONS?
NO? GOOD