Abductive Learning for Neuro-Symbolic Grounded Imitation

Jie-Jing Shao[1*], Hao-Ran Hao[1,2*],
Xiao-Wen Yang[1,2], De-Chuan Zhan[1,2]

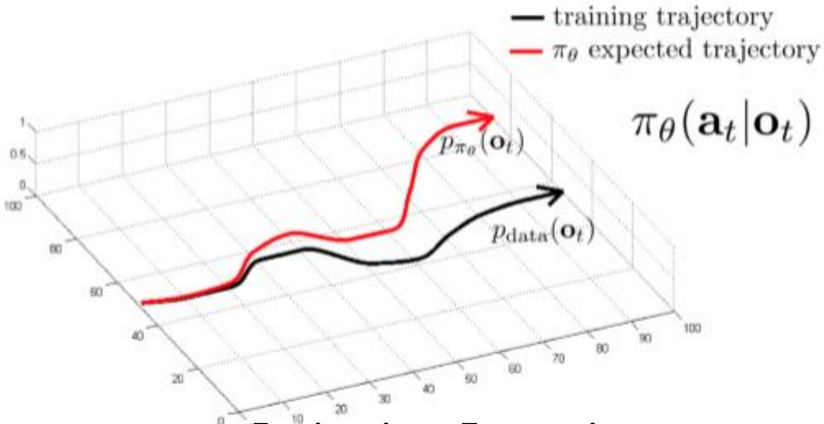
National Key Laboratory for Novel Software Technology,
Nanjing University, China

School of Artificial Intelligence, Nanjing University, China

What is this work about

- Long-Horizon Decision-Making is critical for embodied intelligence.
  - Imitation Learning
  - ✓ Shows promising performance on robotics and auto-driving.
  - ■ Is limited in open environments, especially in the long-horizon tasks.

  - Traditional symbolic planning
  - ✓ Excels at long-horizon tasks via logical reasoning.
  - ■ Typically abstracts away perception with ground-truth symbols, struggles to map visual observations to human-defined symbolic spaces.

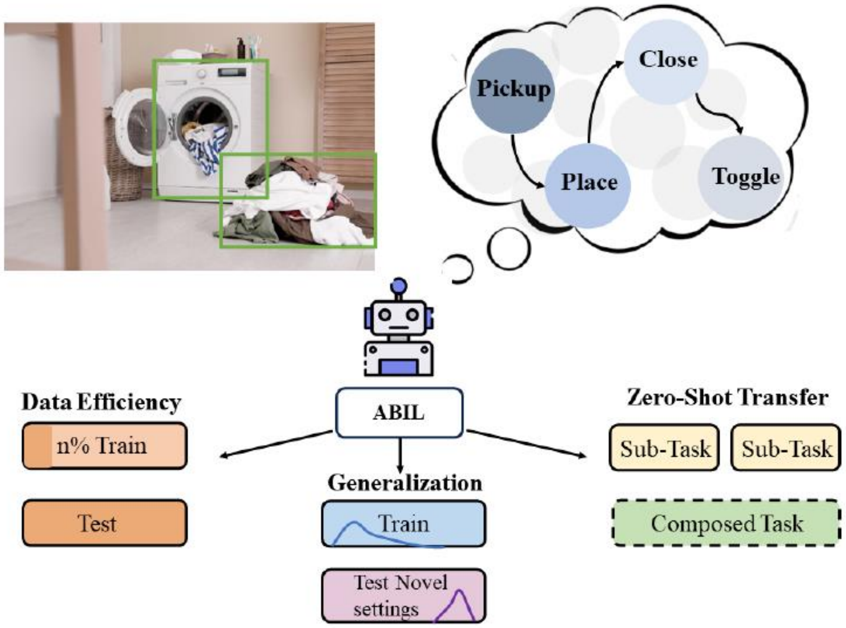


Imitation Learning

Symbolic planning

Such limitations restrict their application in Open environments.

What is this work about



- ✓ In this work, we propose a novel framework Abductive Imitation Learning (ABIL) to combine the benefits of data-driven learning and symbolic-based reasoning.
- ✓ Our **ABIL** shows significantly improved performance on settings of data-efficiency and generalization in the open environments.

Outline

1. **Background**

2. ABIL Framework

3. Empirical Results

4. Conclusion

Long-Horizon Planning

Background

- Previous Studies:

- Imitation learning: is weak at long-horizon tasks
- Symbolic Planning: requires symbolic-level grounding
- Recent efforts on neuro-symbolic solutions[1,2,3]:
  These methods typically assume there are sufficient symbolic information, or only applicable to low-dimensional robotics states.

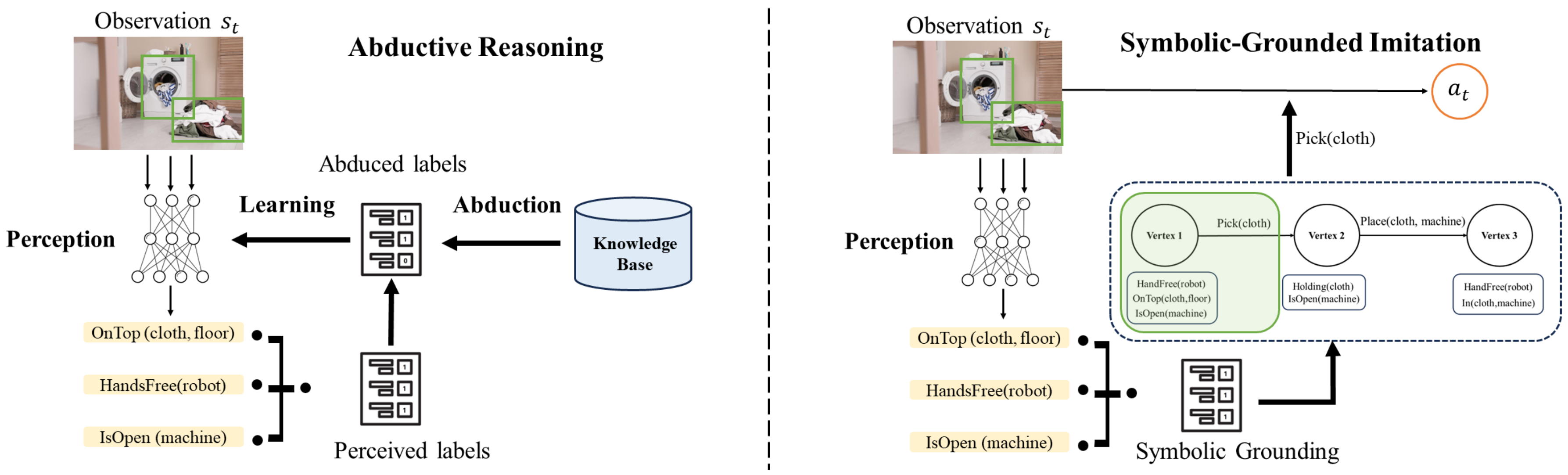
Our Goal

- Help the agent understand demonstrations in symbolic space from high-dimensional visual observations without symbolic-level label.

- Enable long-term logical planning for imitation learning.

[1] Regression Planning Networks. NeurIPS'19
[2] Learning Symbolic Operators for Task and Motion Planning. IROS'21
[3] Programmatically grounded, compositionally generalizable robotic manipulation. ICLR'23

The Overall Framework



Goal:

- Help the agent understand demonstrations in symbolic space from high-dimensional visual observations without symbolic-level label.
- Enable long-term logical planning for imitation learning.

Outline

1. Background & Problem

2. **ABIL Framework**

3. Empirical Results

4. Conclusion

Problem Formulation

Goal-based planning task.

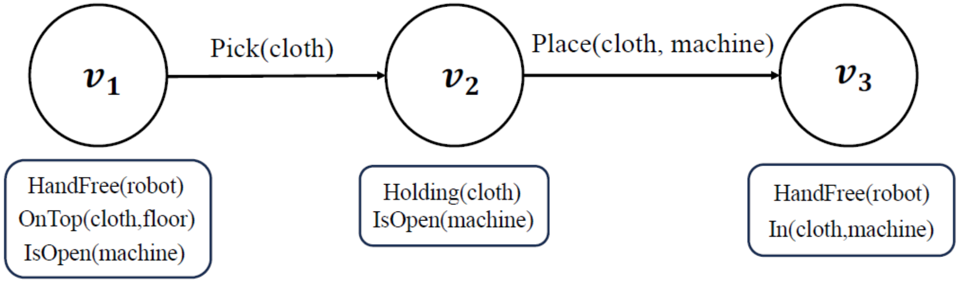
- Environment Definition: $\langle \mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{P}, \mathcal{OP}, \mathcal{S}^0, g \rangle$

  Deterministic, fully-observed environment with object-centric representation.

- Symbolic Knowledge Base:

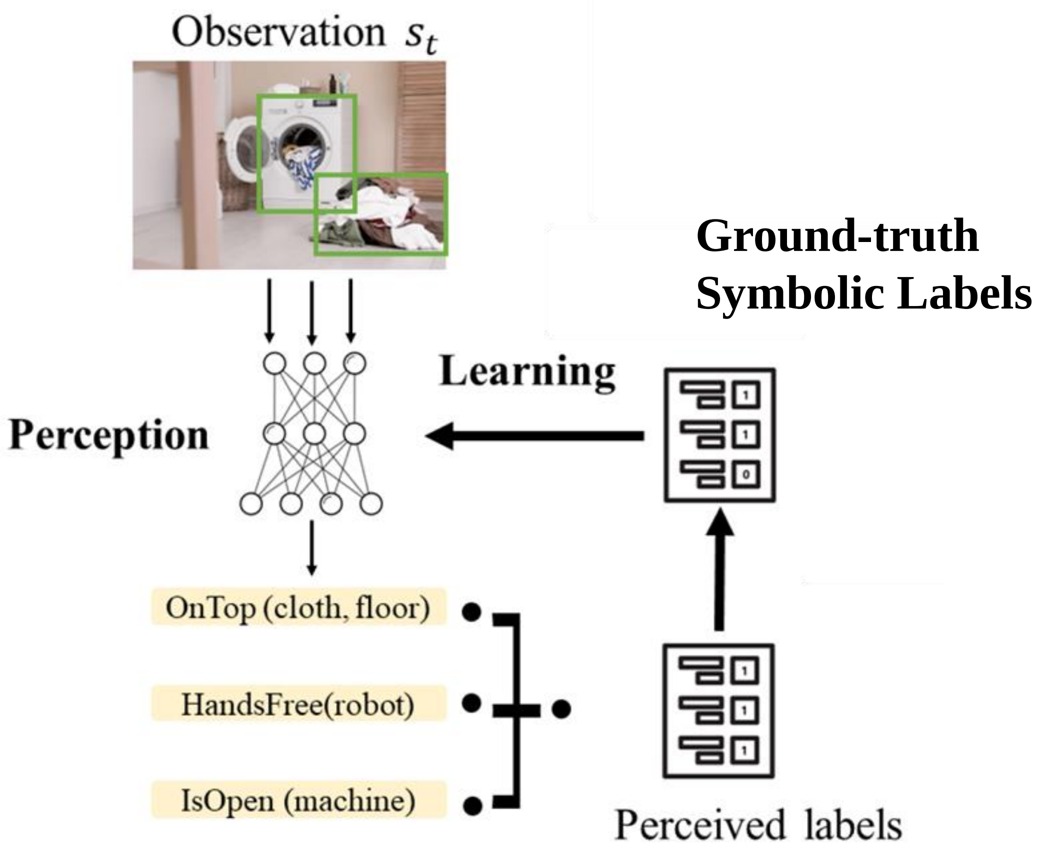
  A finite-state machine, with a directed graph $G = \langle V, E \rangle$

- Each node $v \in V$ contains a set of ground atoms, which can be viewed as the condition of a sub-task.
- Each edge is noted as a tuple $\langle \overline{op}, EFF^+, EFF^- \rangle$.



An example of the knowledge base

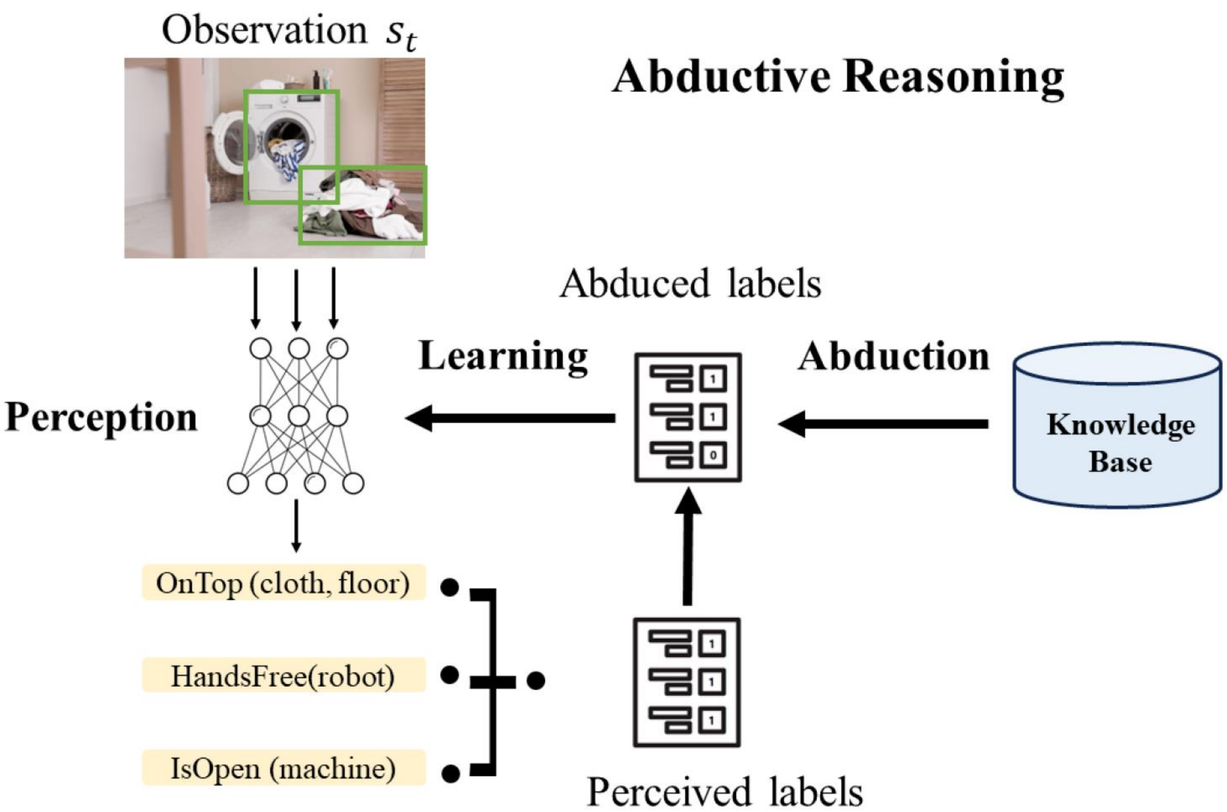
Symbolic-grounded Understanding



A straightforward method: optimize the network with the symbolic labels.

However: Symbolic supervision is typically costly or not available

Symbolic-grounded Understanding



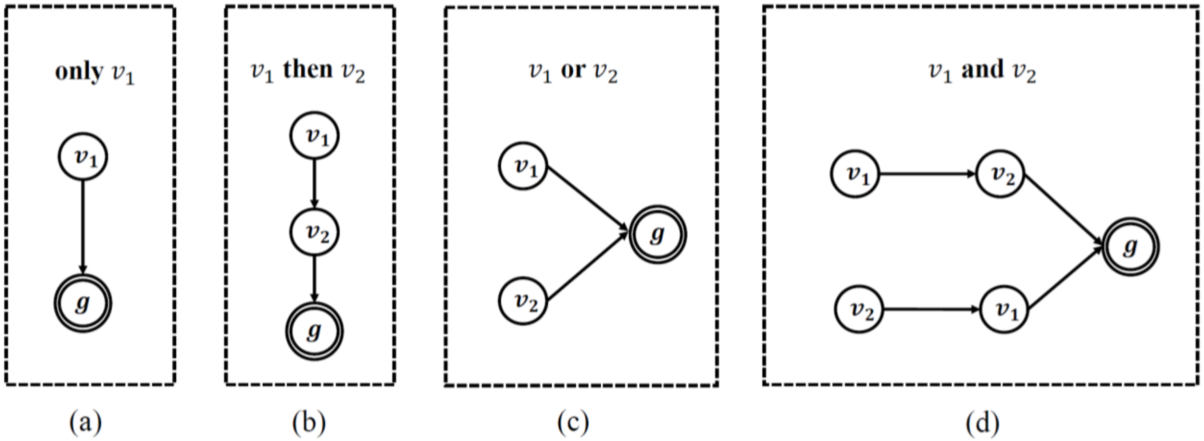
A straightforward method: optimize the network with the symbolic labels.

However: Symbolic supervision is typically costly or not available

We introduce the abductive reasoning to optimize the network.

Abductive Reasoning

➢ Acquire the pseudo label from the knowledge of state machine via abductive reasoning.



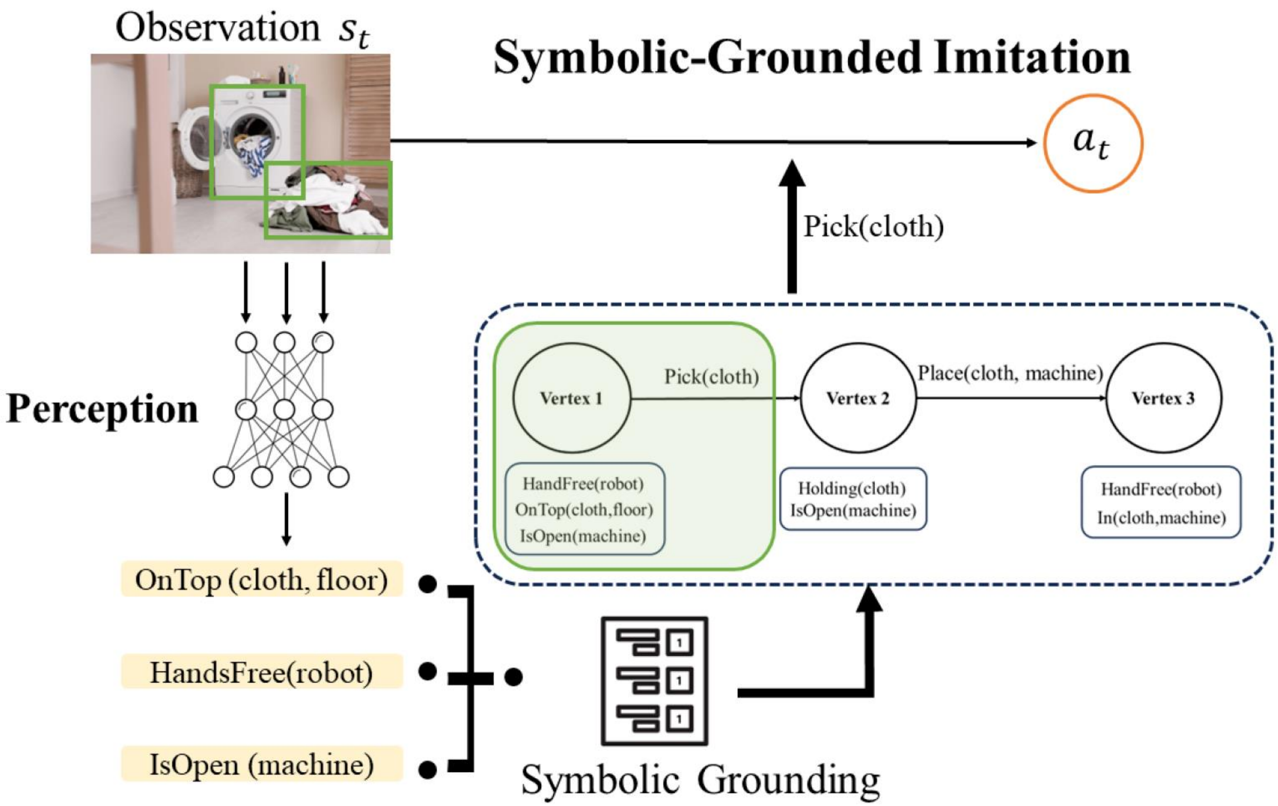
Typical structures of the state machine

- Derive the sequential abduction: $\{z_i^t\}_{t=1}^{T} \vDash G$
- Optimize the perception function f

$$\min_{f} \sum_{s_i \in D} \sum_{t=1}^{T} \mathcal{L}(f(s_i^t), \widehat{z_i^t}),$$

$$\{\widehat{z_i^t}\}_{t=1}^{T} = \arg \min_{\{z_i^t\}_{t=1}^{T}} \sum \|z_i^t - f(s_i^t)\|^2, \quad \text{s.t.} \{z_i^t\}_{t=1}^{T} \vDash G$$

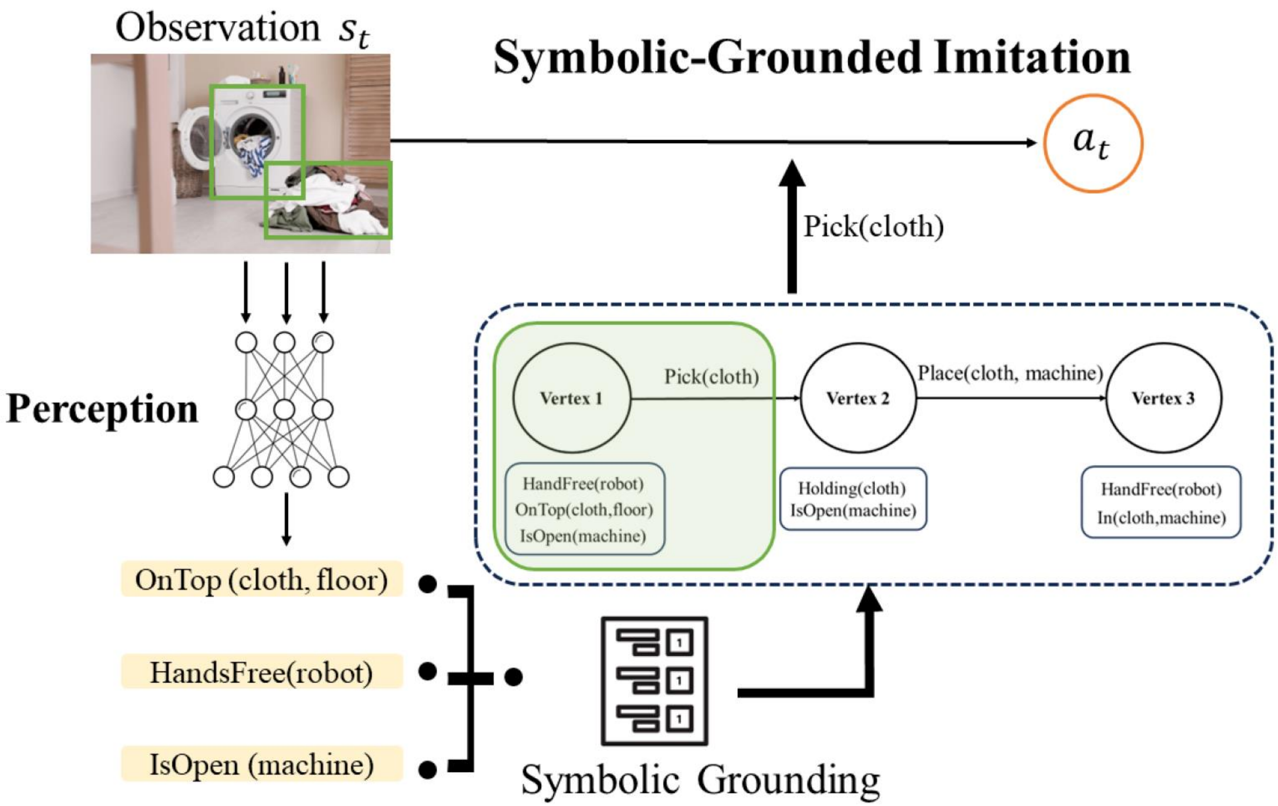
Symbolic-grounded Imitation



- Build the behavioral actor for each logical operator h_{op}, e.g. h_{pick}, h_{place}
- Derive the symbolic states by perception f, and derive the corresponding abstract logical operator

$$\overline{op}^t = \overline{op}_k, \text{s.t.} f(s^t) \models v_k, \exists k \in [0, K)$$

Symbolic-grounded Imitation



- Obtain the desired parameter of the operator $\overline{op}^t$ by reasoning

$$o^t = obj(\overline{op}^t)$$

- Then optimize the behavior actors

$$\min_h \sum_{s_i, a_i \in D} \sum_{t=1}^{T} \mathcal{L}(h_{\overline{op}_i^t}(s_i^t, o^t), a_i^t)$$

ABIL Algorithm

Algorithm 1 Abductive Imitation Learning

Require: Demonstration dataset D, symbolic knowledge G. Number of learning rounds N_R and N_I.

1: **for** $t = 1$ to N_R **do**
2: Get the perceived labels via $f(s)$
3: Get the abduced labels via Eq. 1.
4: Update the perception network f.
5: **end for**
6: **for** $t = 1$ to N_I **do**
7: Get the symbolic states via $f(s)$
8: Get the logical operator $\bar{op}$ via Eq. 2.
9: Update the behavior network $h_{\bar{op}}$ via Eq. 4.
10: **end for**
11: **return** Perception f and behavior $\{h_{\bar{op}}\}, \bar{op} \in \mathcal{OP}$.

- A two-stage learning algorithm.
- Embed high-level logical reasoning into the imitation learning process.

Outline

1. Background & Problem

2. ABIL Framework

3. Empirical Results

4. Conclusion

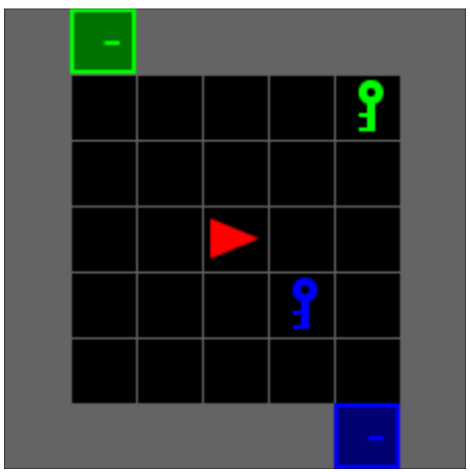
Setup

Three diverse environments

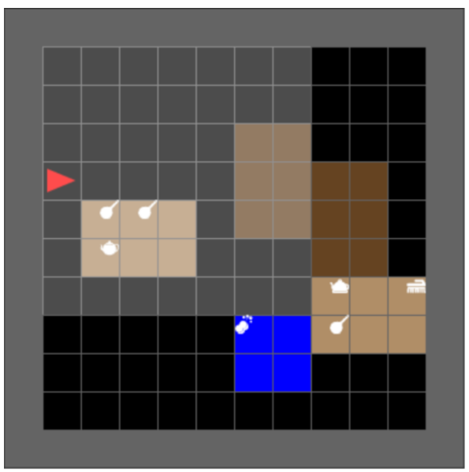
- BabyAI
  - ✓ *Learning with logical instruction*
- Mini-BEHAVIOR
  - ✓ *Household Agent*
- CLIPort
  - ✓ *Robotic manipulation*

Baseline Methods

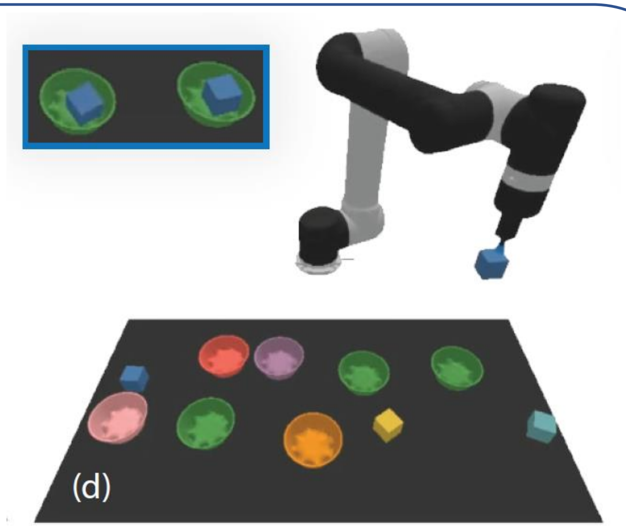
- Behavior Cloning (BC)
- Decision Transformer (DT)
- PDSketch



BabyAI



Mini-BEHAVIOR



CLIPort

Evaluation on BabyAI

Task	Eval	BC	DT	PDSketch	ABIL-BC	ABIL-DT
GotoSingle	Basic	**1.00**	0.893±0.049	**1.00**	**1.00**	0.900±0.036
Goto	Basic	0.843±0.006	0.720±0.044	**1.00**	0.900±0.046	0.853±0.038
	Gen	0.743±0.045	0.583±0.049	**1.00**	0.777±0.032	0.793±0.029
Pickup	Basic	0.723±0.031	0.490±0.040	**0.990±0.010**	0.847±0.025	0.845±0.035
	Gen	0.533±0.031	0.320±0.070	**0.973±0.012**	0.730±0.010	0.763±0.051
Open	Basic	0.933±0.025	0.493±0.059	**1.00**	0.963±0.021	0.903±0.064
	Gen	0.877±0.015	0.440±0.078	**1.00**	0.927±0.032	0.813±0.064
Put	Basic	**0.950±0.044**	0.910±0.036	0.650±0.026	0.930±0.010	0.920±0.026
	Gen	0.037±0.012	0.207±0.092	0.560±0.052	**0.917±0.015**	0.877±0.025
Unlock	Basic	0.957±0.012	0.885±0.035	0.293±0.051	0.967±0.023	**0.993±0.012**
	Gen	0.910±0.030	0.883±0.075	0.247±0.051	0.963±0.006	**0.993±0.012**
Averaged time per evaluation		0.174 seconds	0.260 seconds	8.17 seconds	0.320 seconds	0.354 seconds

ABIL effectively improves the performance of imitation learning methods.

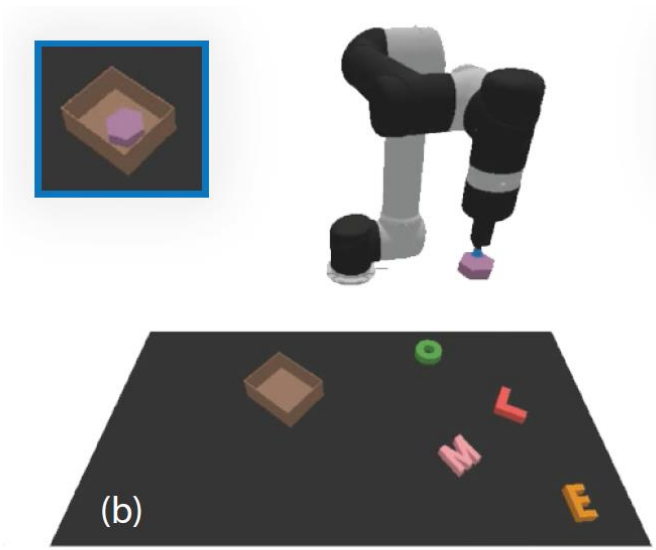
Results on Mini-BEHAVIOR

Task	Eval	BC	DT	PDSketch	ABIL-BC	ABIL-DT
Boxing books up	Basic	0.707±0.041	**0.713±0.035**	> 5 minutes	0.709±0.077	0.661±0.094
	Gen	0.335±0.177	0.519±0.191		**0.644±0.172**	0.625±0.087
Cleaning A Car	Basic	0.417±0.047	0.313±0.091	> 5 minutes	**0.423±0.032**	0.330±0.050
	Gen	0.170±0.036	0.147±0.083		**0.253±0.047**	0.170±0.078
Cleaning shoes	Basic	0.482±0.086	0.427±0.042	> 5 minutes	**0.598±0.068**	0.478±0.020
	Gen	0.030±0.005	0.053±0.046		**0.390±0.102**	0.290±0.026
Collect misplaced items	Basic	0.460±0.030	0.299±0.015	> 5 minutes	**0.617±0.061**	0.457±0.007
	Gen	0.325±0.074	0.261±0.023		**0.423±0.051**	0.387±0.028
Installing a printer	Basic	0.903±0.023	0.927±0.021	0.343±0.032	0.887±0.021	**0.937±0.023**
	Gen	0.003±0.006	0.300±0.147	0.310±0.046	0.727±0.047	**0.757±0.107**
Laying wood floors	Basic	0.616±0.062	0.638±0.027	> 5 minutes	**0.644±0.043**	0.643±0.031
	Gen	0.068±0.018	0.366±0.041		**0.628±0.057**	0.374±0.040
Making tea	Basic	0.607±0.015	0.583±0.105	> 5 minutes	**0.687±0.038**	0.607±0.029
	Gen	0.070±0.078	0.113±0.105		0.370±0.131	**0.493±0.124**
Moving boxes to storage	Basic	0.753±0.083	0.780±0.017	> 5 minutes	0.767±0.012	**0.787±0.032**
	Gen	0.417±0.110	0.617±0.042		**0.730±0.017**	0.673±0.119
Opening packages	Basic	0.947±0.045	0.963±0.034	0.020±0.010	0.978±0.010	**0.990±0.009**
	Gen	0.295±0.180	0.548±0.065	0.020±0.010	0.905±0.018	**0.918±0.033**
Organizing file cabinet	Basic	0.156±0.047	0.522±0.067	> 5 minutes	0.231±0.021	**0.562±0.037**
	Gen	0.083±0.012	0.382±0.112		0.095±0.009	**0.454±0.074**
Putting away dishes	Basic	0.811±0.031	0.828±0.052	> 5 minutes	**0.883±0.043**	0.813±0.022
	Gen	0.141±0.111	0.547±0.296		**0.830±0.013**	0.739±0.072
Sorting books	Basic	0.601±0.032	0.543±0.053	> 5 minutes	0.618±0.012	**0.631±0.055**
	Gen	0.131±0.047	0.220±0.010		0.338±0.078	**0.412±0.038**
Throwing away leftovers	Basic	0.833±0.080	0.890±0.029	> 5 minutes	**0.924±0.014**	0.888±0.039
	Gen	0.222±0.167	0.653±0.039		0.713±0.069	**0.729±0.031**
Washing pots and pans	Basic	0.342±0.022	0.227±0.079	> 5 minutes	**0.349±0.063**	0.184±0.024
	Gen	0.099±0.168	0.028±0.016		**0.242±0.110**	0.153±0.024
Watering houseplants	Basic	0.814±0.034	0.806±0.020	> 5 minutes	**0.843±0.010**	0.835±0.022
	Gen	0.002±0.004	0.187±0.113		0.545±0.151	**0.734±0.063**
Averaged time per evaluation		1.48 seconds	2.09 seconds	> 5 minutes	2.88 seconds	2.98 seconds

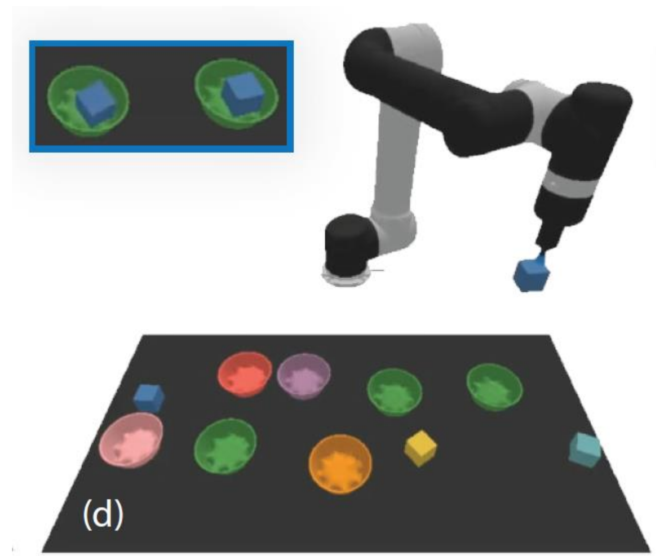
ABIL demonstrates great performance under the open enviornments.

Results on CLIPort

Task	BC	DT	ABIL-BC	ABIL-DT
Packing-5shapes	0.580±0.252	0.607±0.223	**0.983±0.015**	0.903±0.085
Packing-20shapes	0.207±0.006	0.180±0.026	**0.940±0.030**	0.857±0.025
Put-4blocks-in-5bowl	0.365±0.141	0.319±0.068	**0.962±0.012**	0.917±0.033



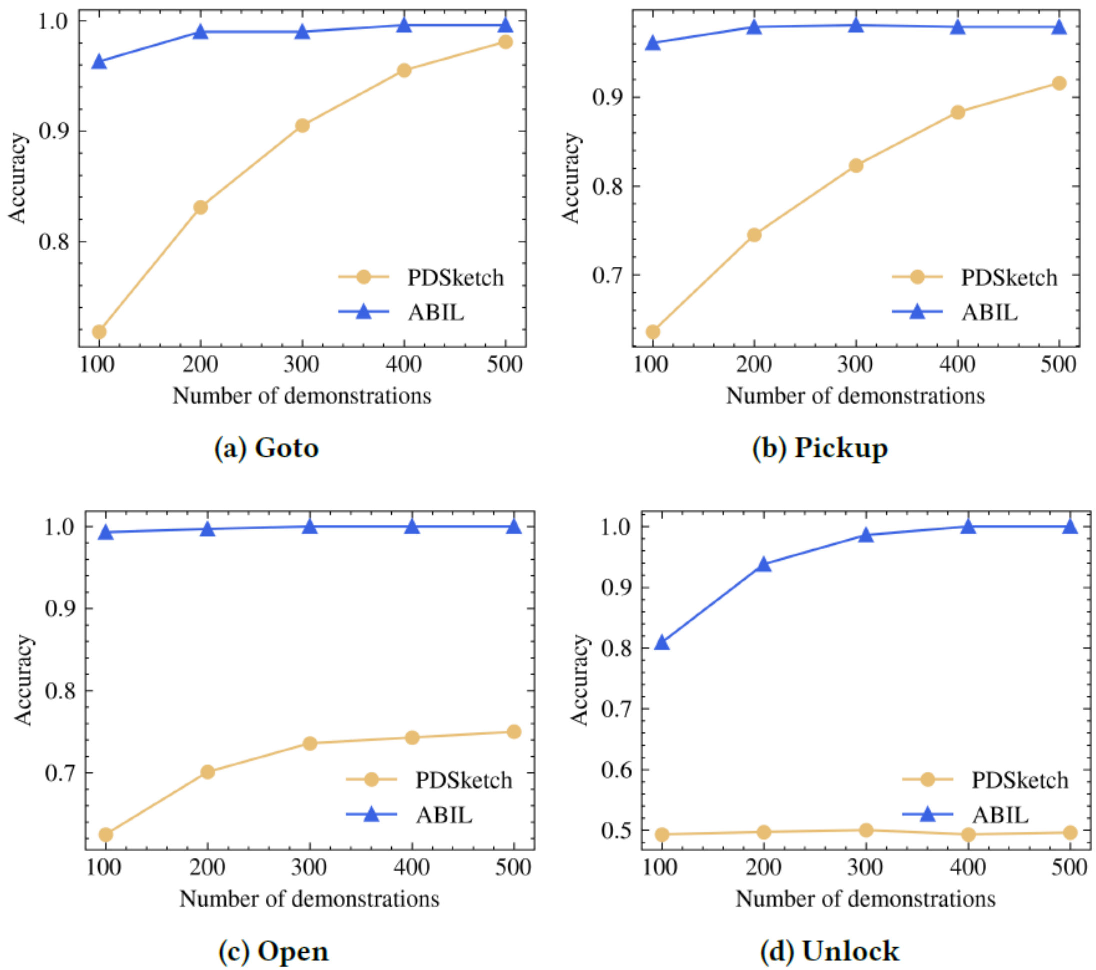
Packing-shapes



Put-blocks-in-bowls

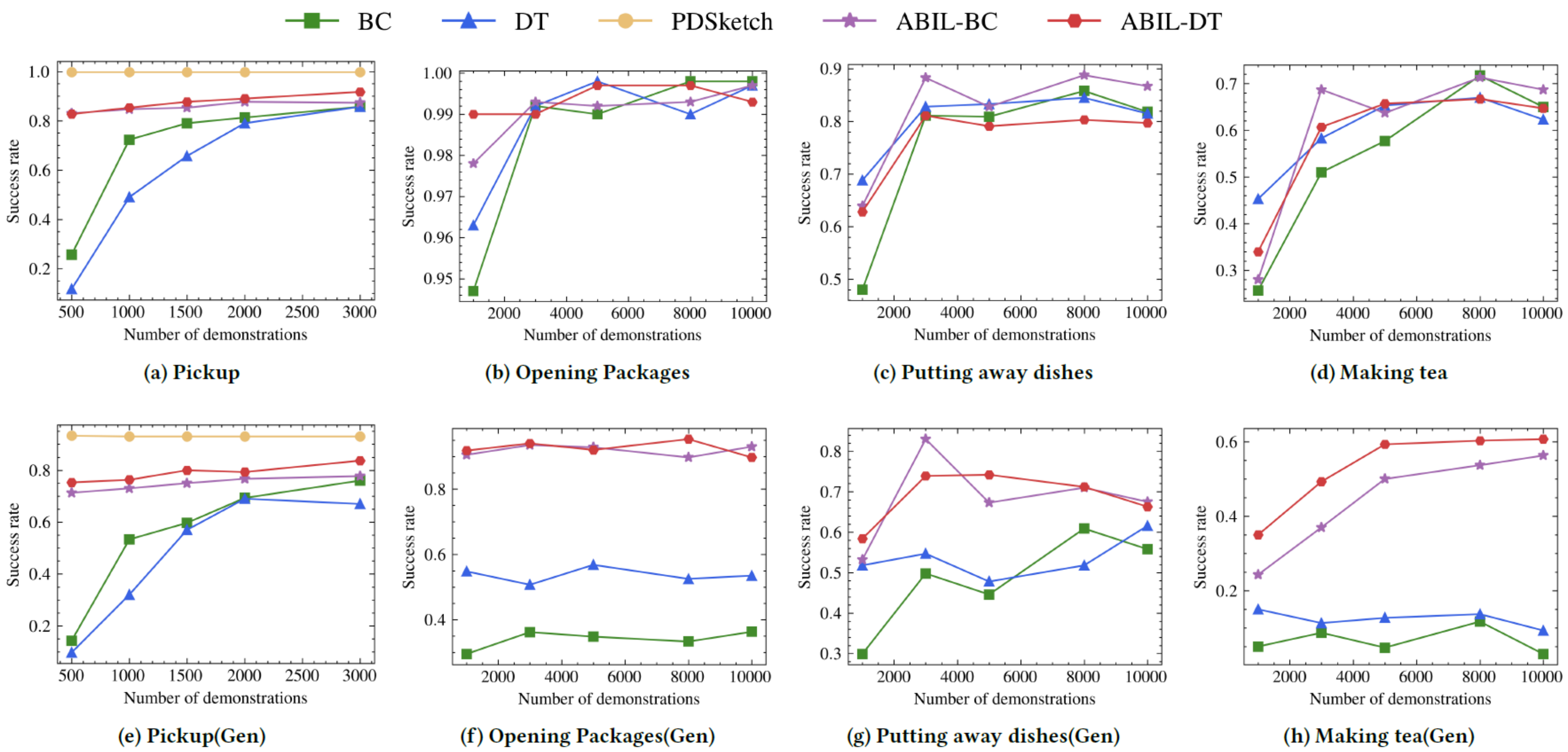
ABIL gives outstanding results in CLIPort Environment.

Comparison of Neural-Symbolic Grounding



ABIL outperforms in **understanding the environment** accurately.

Data Efficiency and Generalization



ABIL improves the **data efficiency** of the BC and DT baselines, achieves significant **generalization improvement** in the out-of-distribution evaluation

Compositional Generalization

Domain	BabyAI		
	Train		Eval
Task	Pickup	Open	Unlock
BC	0.760±0.056	0.983±0.021	0.120±0.010
DT	0.783±0.031	0.957±0.031	0.057±0.051
PDSketch	**0.970±0.010**	0.990±0.010	0.127±0.021
ABIL-BC	0.937±0.021	**1.00**	0.980±0.026
ABIL-DT	0.925±0.007	**1.00**	**0.993±0.012**

Domain	Mini-BEHAVIOR					
	Train	Eval		Train	Eval	
Task	Open 1	Open 2	Open 3	Throw 1	Throw 2	Throw 3
BC	0.950±0.087	0.012±0.010	0.002±0.004	0.703±0.085	0.117±0.070	0.053±0.045
DT	**1.00**	0.037±0.025	0.024±0.008	0.770±0.026	0.182±0.008	0.056±0.003
PDSketch	0.467±0.057	0.020±0.010	> 5 minutes	0.013±0.006	> 5 minutes	> 5 minutes
ABIL-BC	0.997±0.006	0.818±0.014	0.551±0.032	0.763±0.049	0.638±0.052	0.536±0.082
ABIL-DT	**1.00**	**0.840±0.035**	**0.631±0.041**	**0.803±0.051**	**0.650±0.049**	**0.585±0.120**

ABIL has the ability to **zero-shot generalize** to novel composed tasks.

Outline

1. Background & Problem

2. ABIL Framework

3. Empirical Results

4. Conclusion

Conclusion

- ➢ In this paper, we propose a novel framework: ABIL
- ✓ A novel framework which combines the benefits of data-driven learning and symbolic-based reasoning.
- ✓ Extensive experiments demonstrate the effectiveness and generality of ABIL.

Future work

- ➢ Learning with accurate and incomplete knowledge base

Thank you!

If you are interested in, feel free to contact us:

Hao-Ran Hao (hhr277133291@gmail.com)

Jie-Jing Shao (shaojj@lamda.nju.edu.cn)