Visual Imitation Learning

Generalization, Perceptual Grounding, and Abstraction

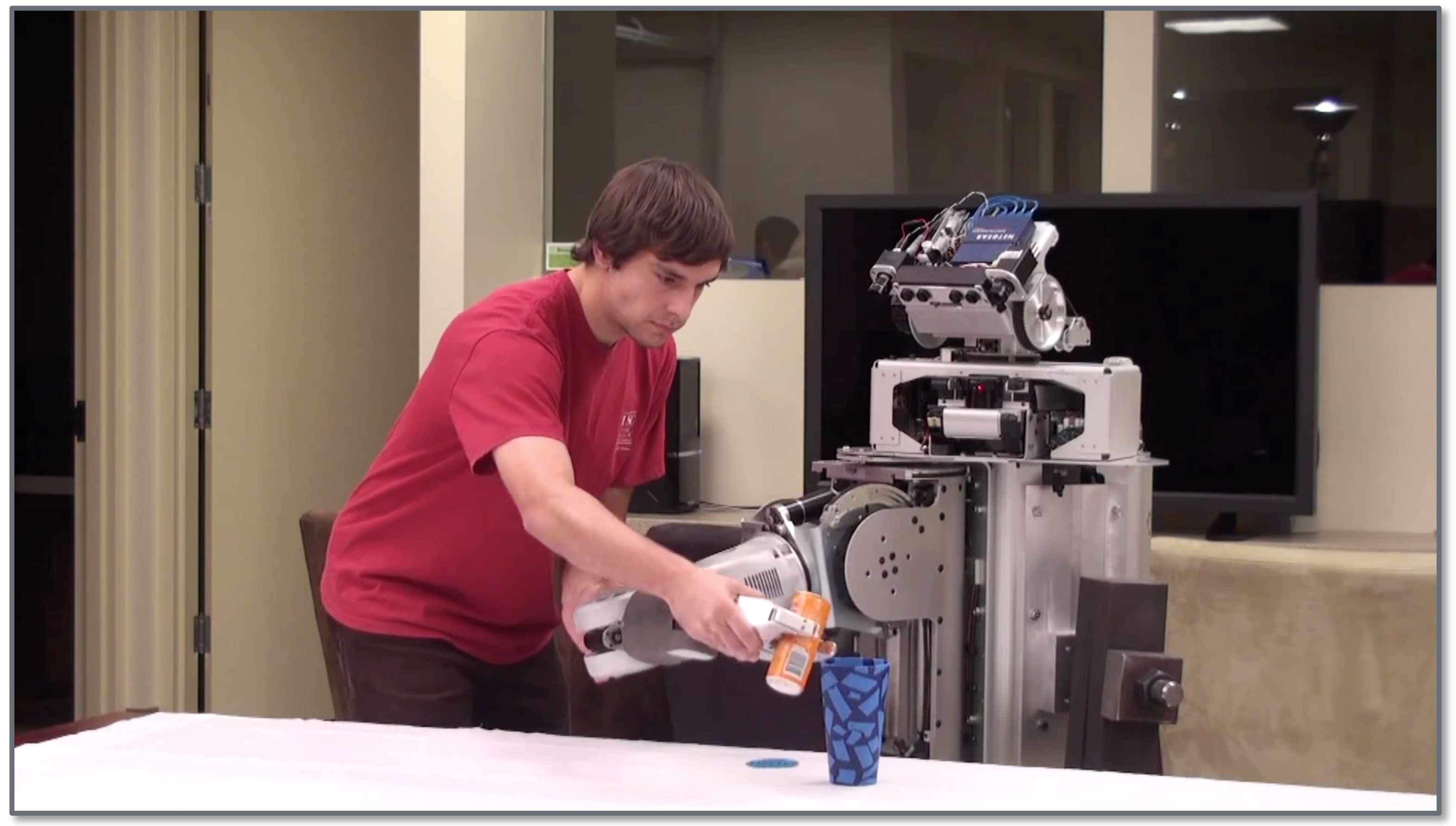
Yuke Zhu

RSS 2020

TEXAS
The University of Texas at Austin

NVIDIA

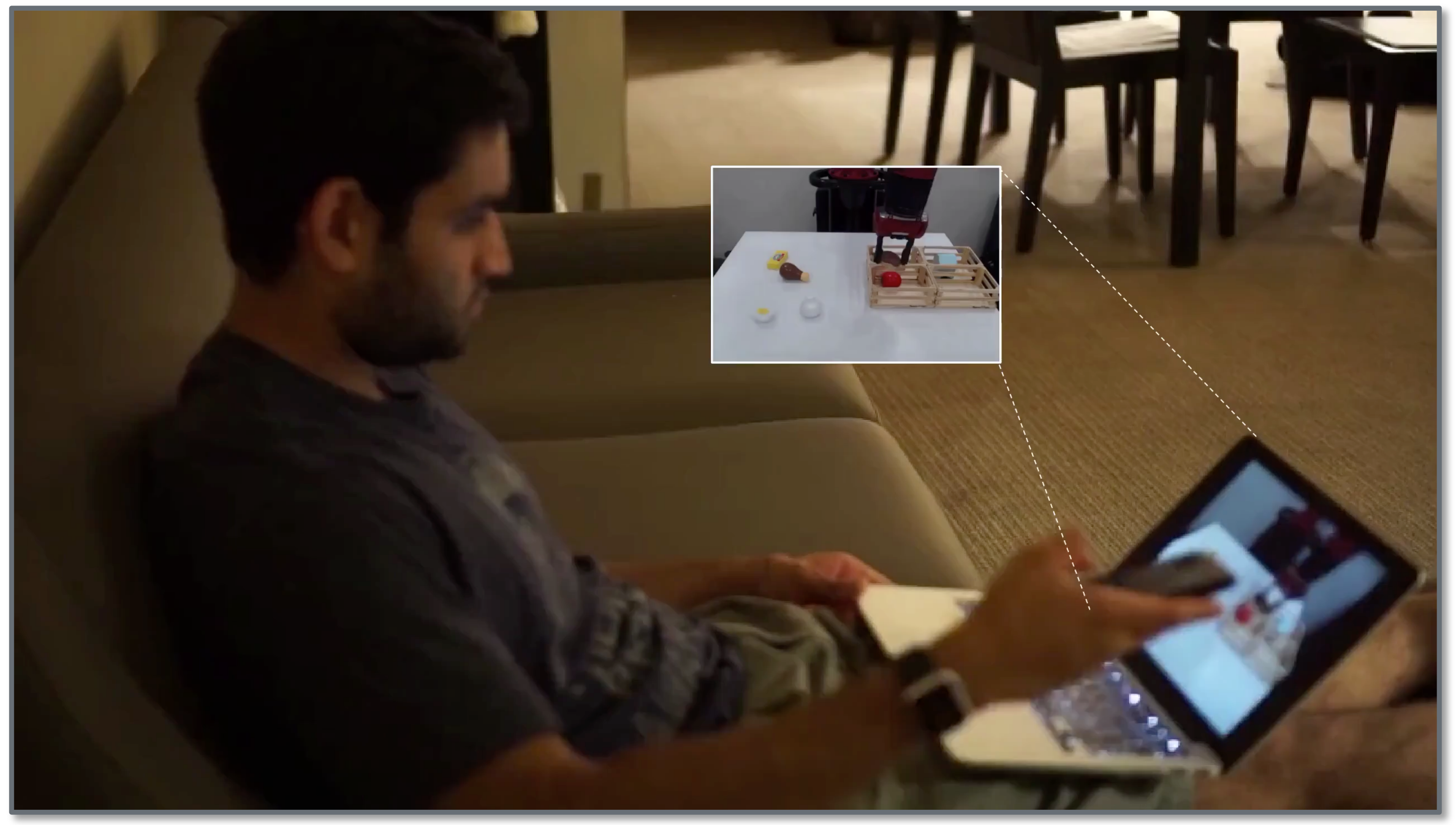
Imitation Learning in Robotics



Kinesthetic Teaching

"Dynamic Movement Primitives"

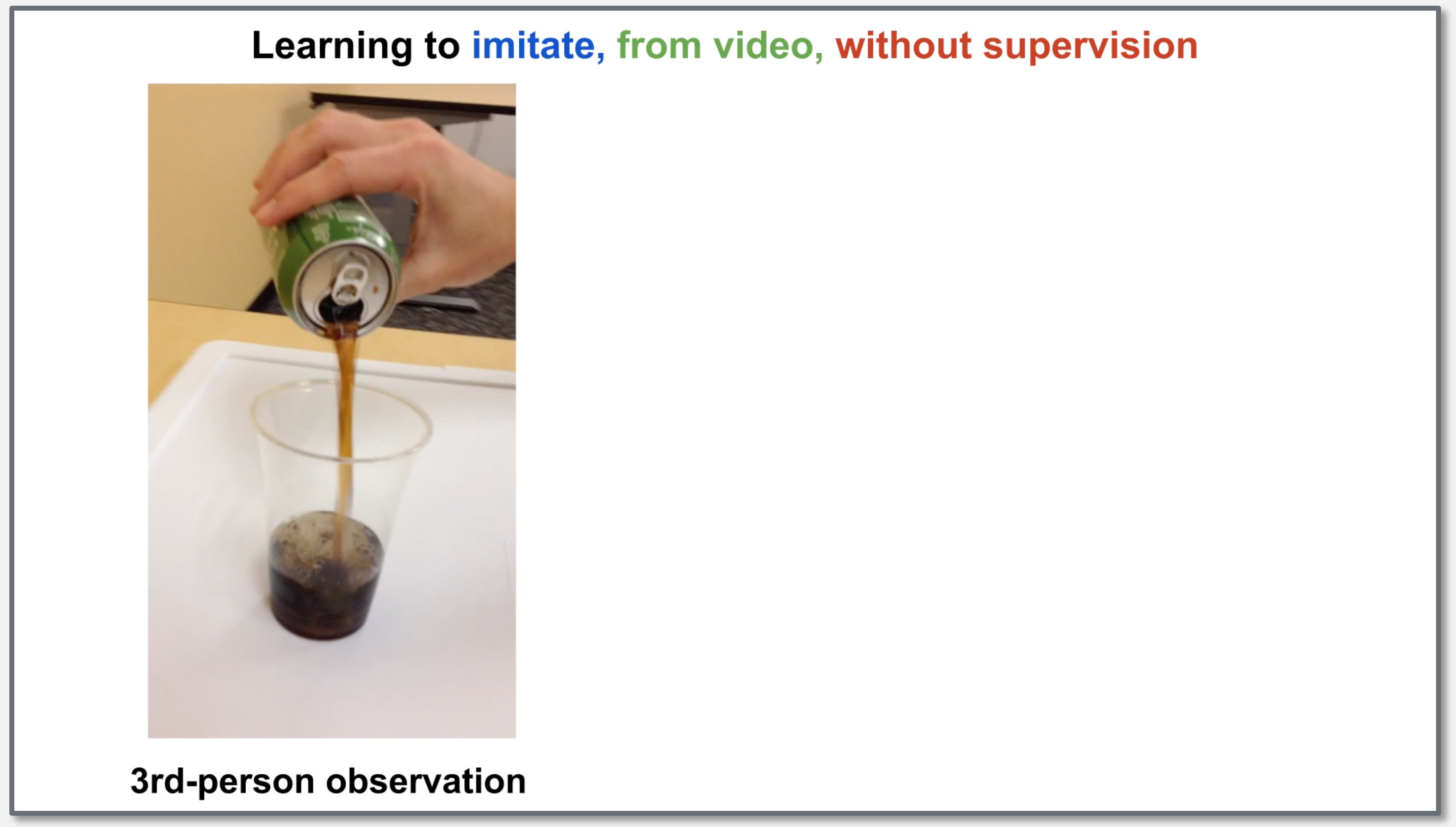
[Schaal et al. 2002; Pastor et al. 2009]



Teleoperation

"RoboTurk"

[Mandlekar et al. IROS'19; Mandlekar CoRL'18]



Imitation from Observation

"Time Contrastive Network"

[Sermanet et al. ICRA 2018]

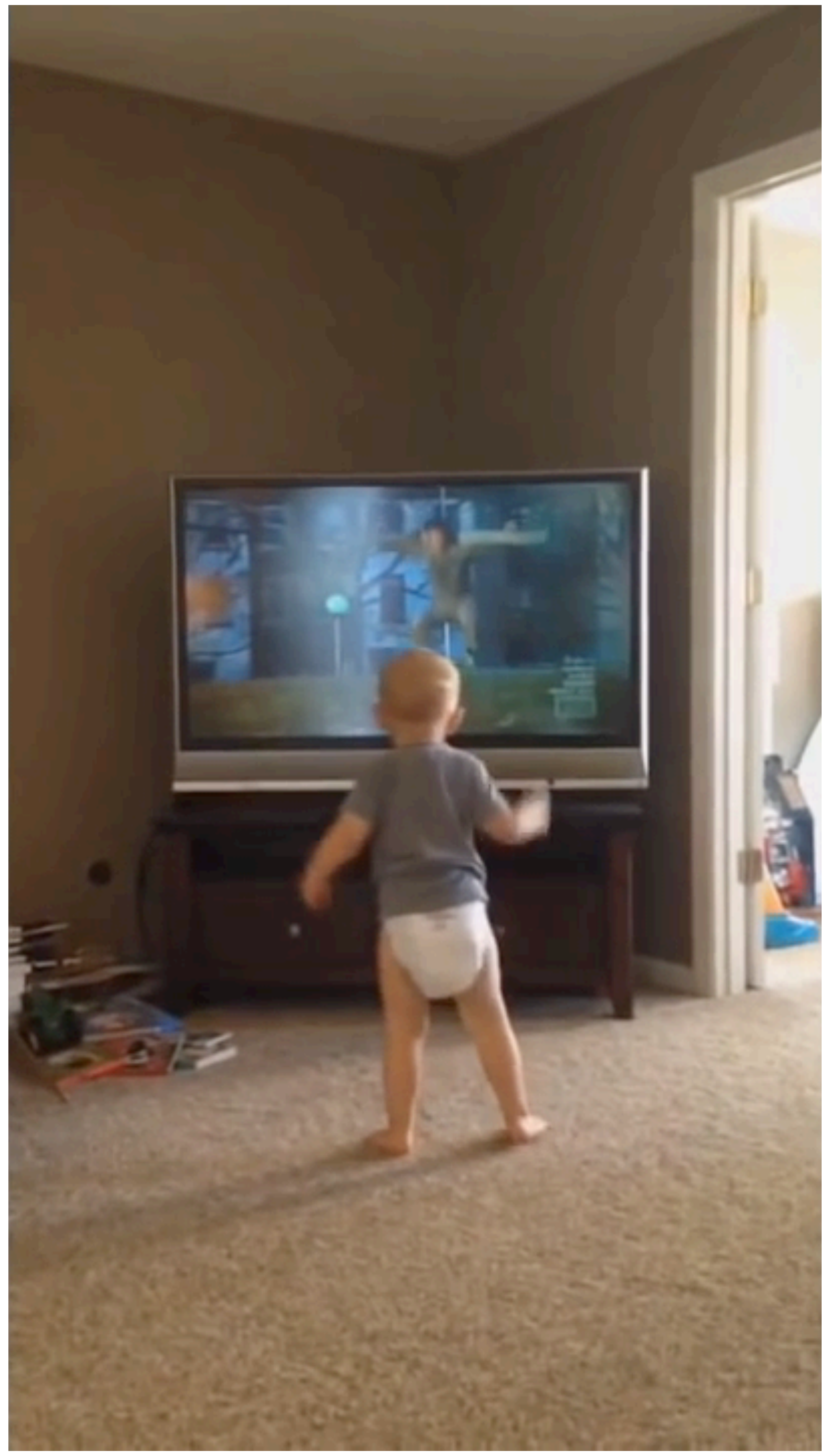
Why Imitation from Observation?

Humans learn efficiently from **visual demonstrations**.

Imitation of Televised Models by Infants

Andrew N. Meltzoff, *Child Development* 1988

Babies (14-24 months) can learn by imitating demonstrations from the TV screen.



Meltzoff & Moore 1977; Meltzoff & Moore 1989, Meltzoff 1988

Why Imitation from Observation?

Humans learn efficiently from **visual demonstrations**.

THE VERGE TECH SCIENCE CULTURE CARS REVIEWS LONGFORM VIDEO MORE

TECH \ YOUTUBE \ CULTURE

Half of YouTube viewers use it to learn how to do things they've never done

10

Some of us are on there just to pass the time, though

By Patricia Hernandez | @xpatriciah | Nov 7, 2018, 12:36pm EST

SHARE

how to

how to **make slime**
how to **tie a tie**
how to **draw**
how to **basic**
how to **get boogie down dance**
how to **cake it**
how to **train your dragon 3**
how to **get the galaxy skin in fortnite**
how to **make slime without glue**
how to **solve a rubik's cube**

Report search predictions

NOVEMBER 7, 2018

Many Turn to YouTube for Children's Content, News, How-To Lessons

An analysis of videos suggested by the site's recommendation engine finds that users are directed toward progressively longer and more popular content

BY AARON SMITH, SKYE TOOR AND PATRICK VAN KESSEL

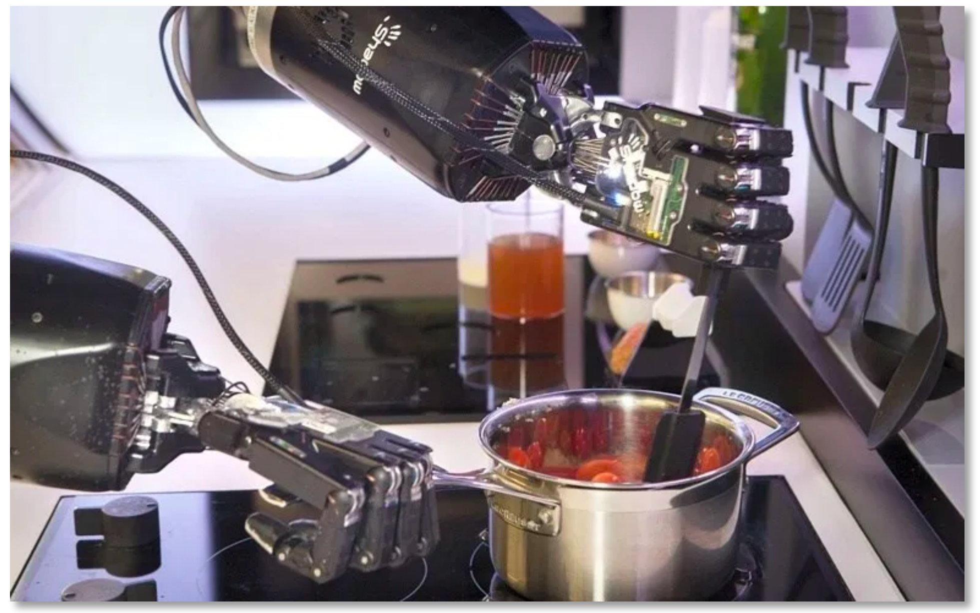
(MaaHoo Studio/Getty Images)

Why Imitation from Observation?

Towards **large-scale imitation learning** in the wild

- **Demonstrator flexibility**: Allow humans to perform the task using their own body.
- **High-DoF robots**: Get around the difficulty of controlling complex robot morphologies with high degrees of freedom.
- **Massive video data sources**: Internet videos of human doing tasks -- enabling "web-scale" imitation learning

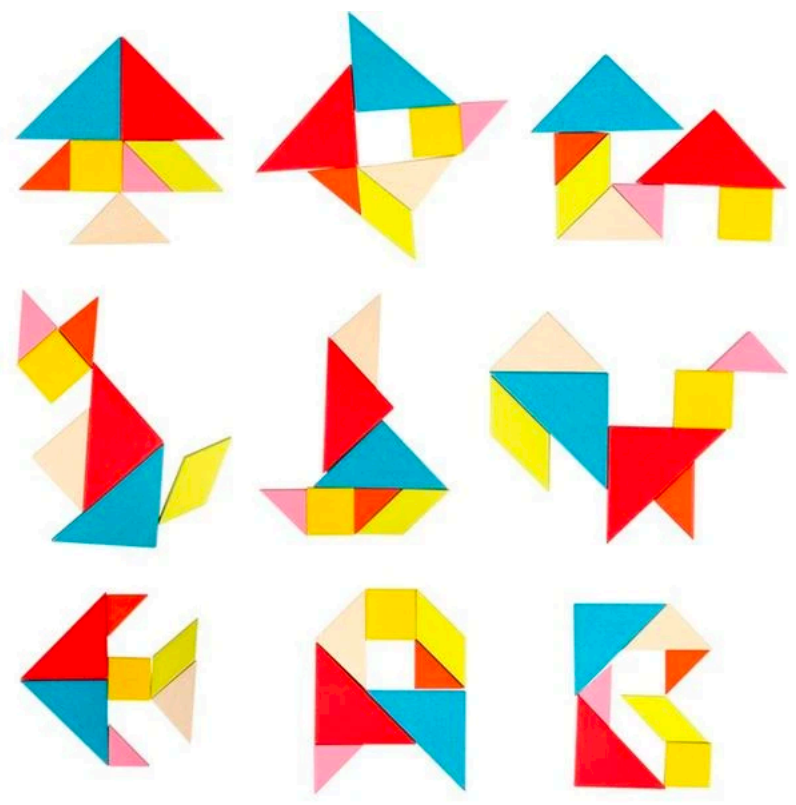
video demonstration

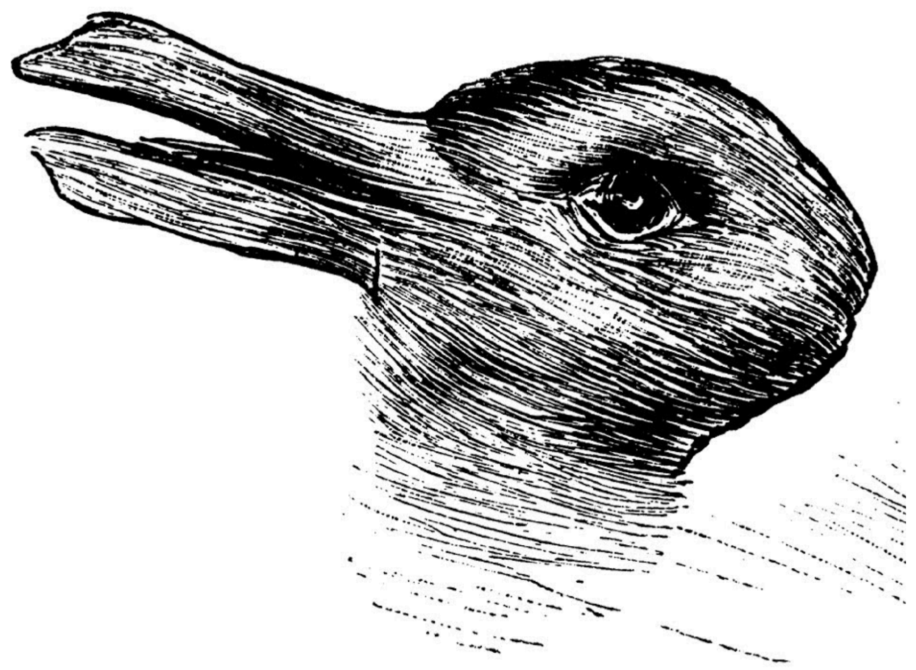


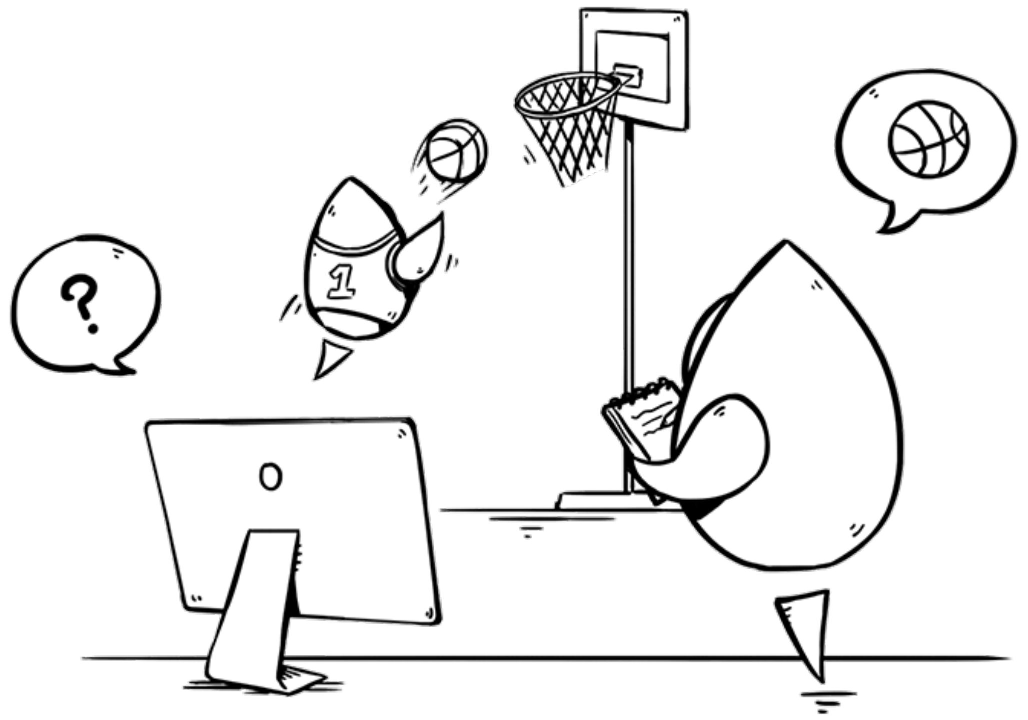
robot execution

Today’s Talk

Visual imitation Learning from video demonstrations







Compositional Generalization

How can we generalize across task structures and task goals?

Perceptual Uncertainty

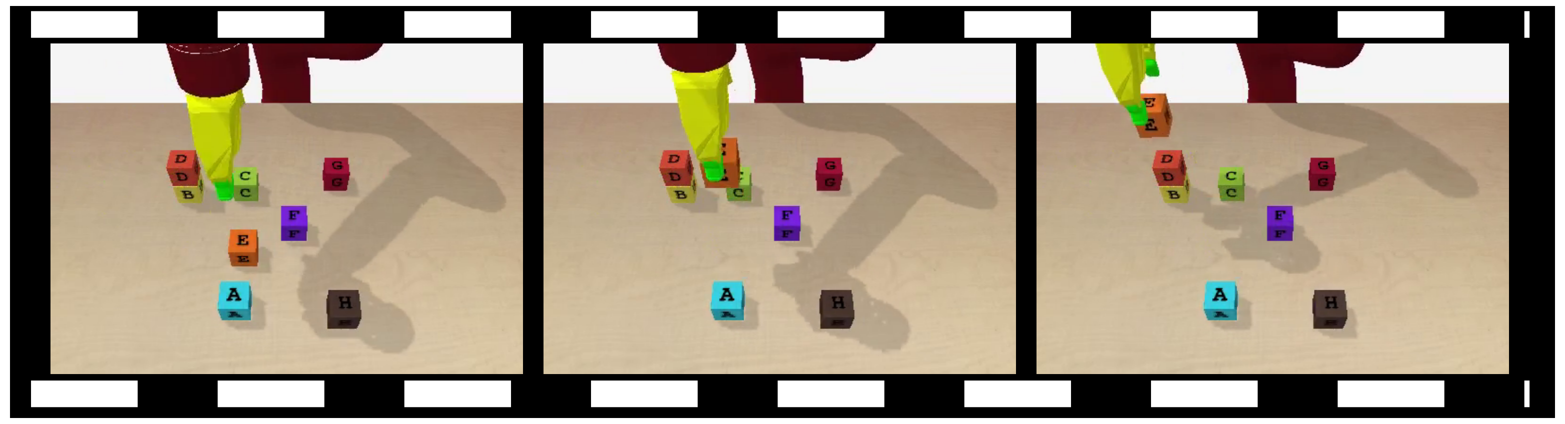
How to address perceptual uncertainty arising from visual imitation?

Long-horizon Tasks

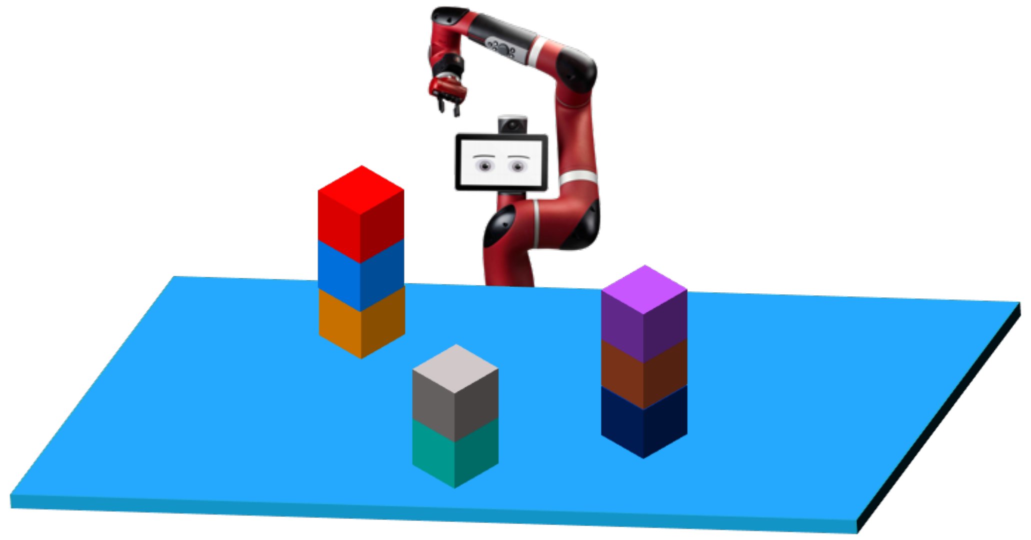
How can we extrapolate to long-horizon tasks?

Visual Imitation Learning

one-shot visual imitation learning as **meta-learning**



meta-policy



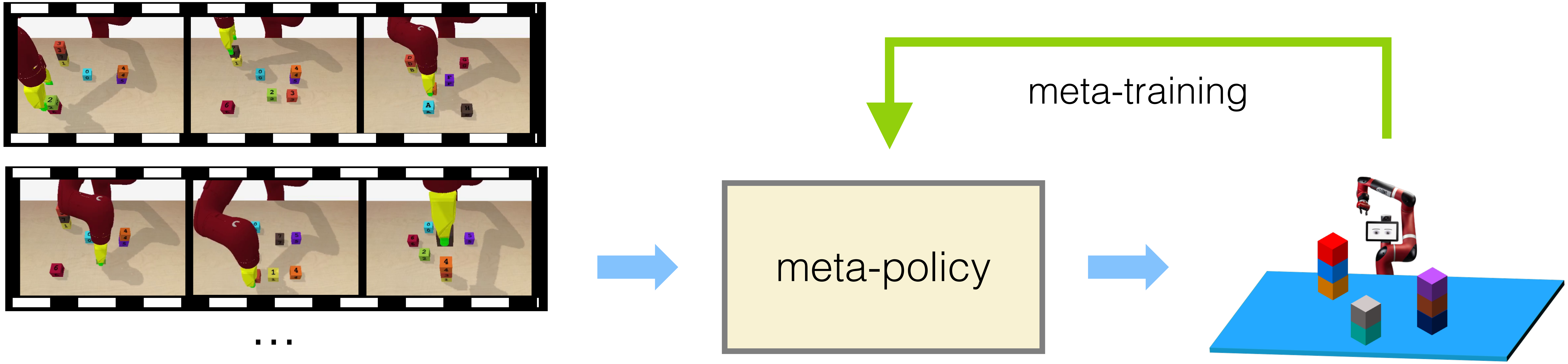
single video demonstration

policy for the demonstrated task

[Xu*, Nair*, et al. ICRA 2018; Huang*, Nair*, Xu*, et al. CVPR'2019; Huang et al. IROS'19]

Visual Imitation Learning

one-shot visual imitation learning as **meta-learning**



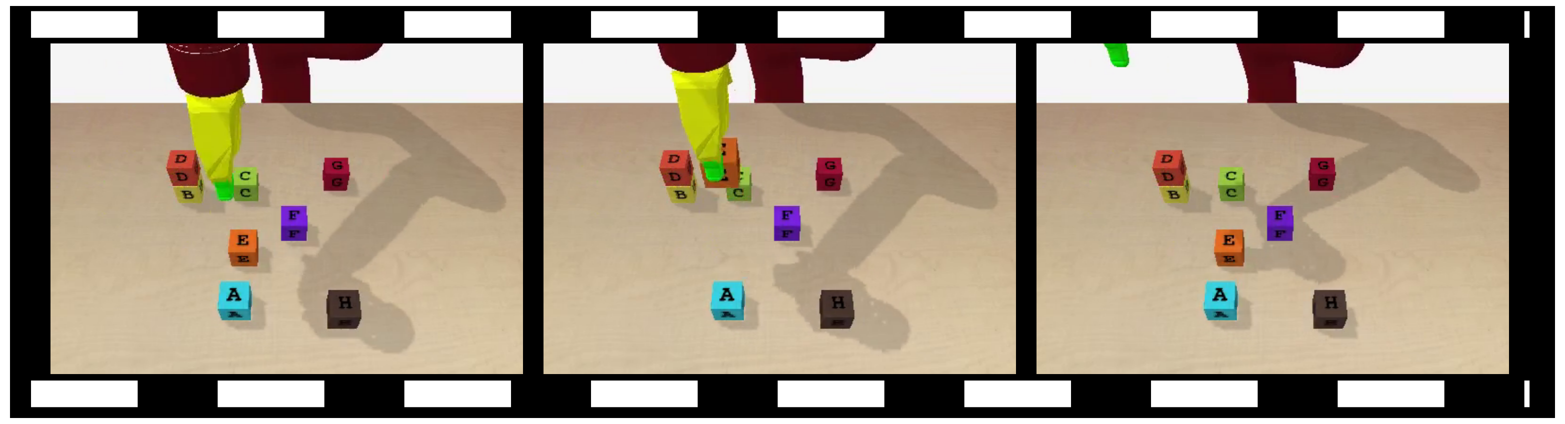
training videos (seen tasks)

policy for the demonstrated task

[Xu*, Nair*, et al. ICRA 2018; Huang*, Nair*, Xu*, et al. CVPR'2019; Huang et al. IROS'19]

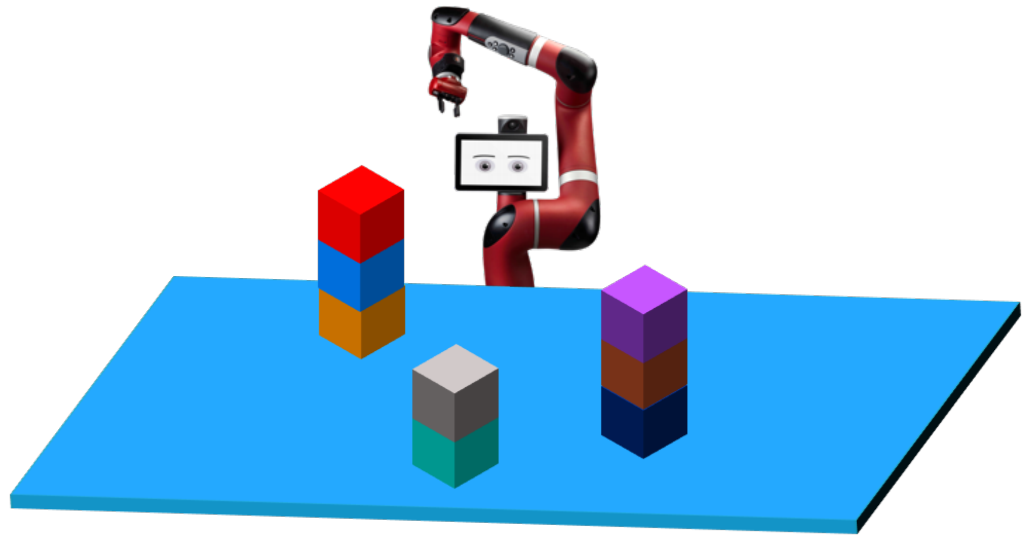
Visual Imitation Learning

one-shot visual imitation learning as **meta-learning**



single test video
(unseen task)

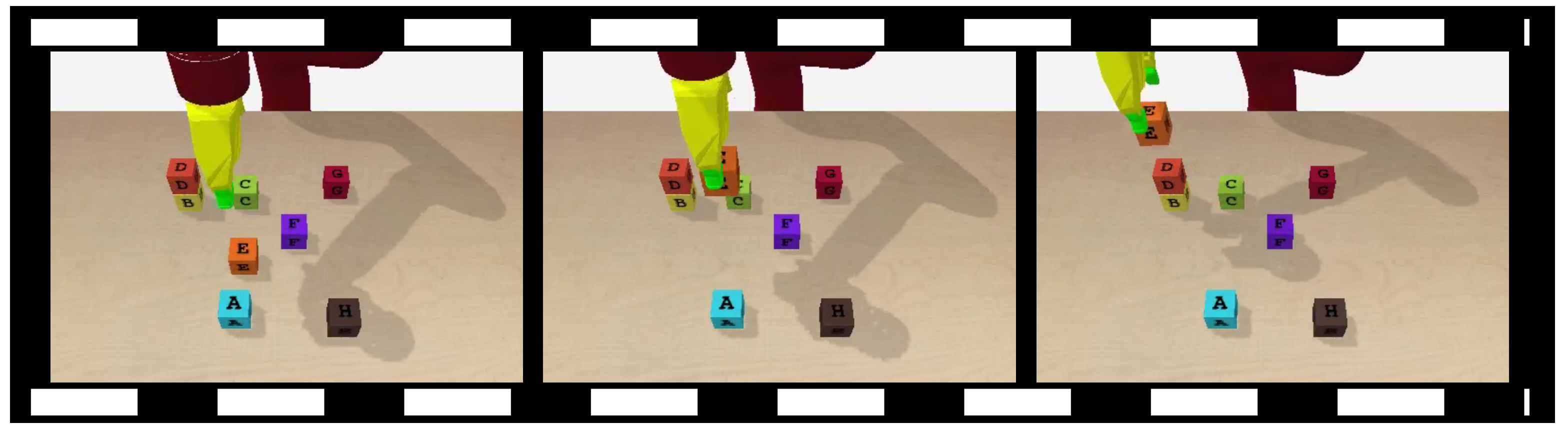
meta-policy



policy for the
demonstrated task

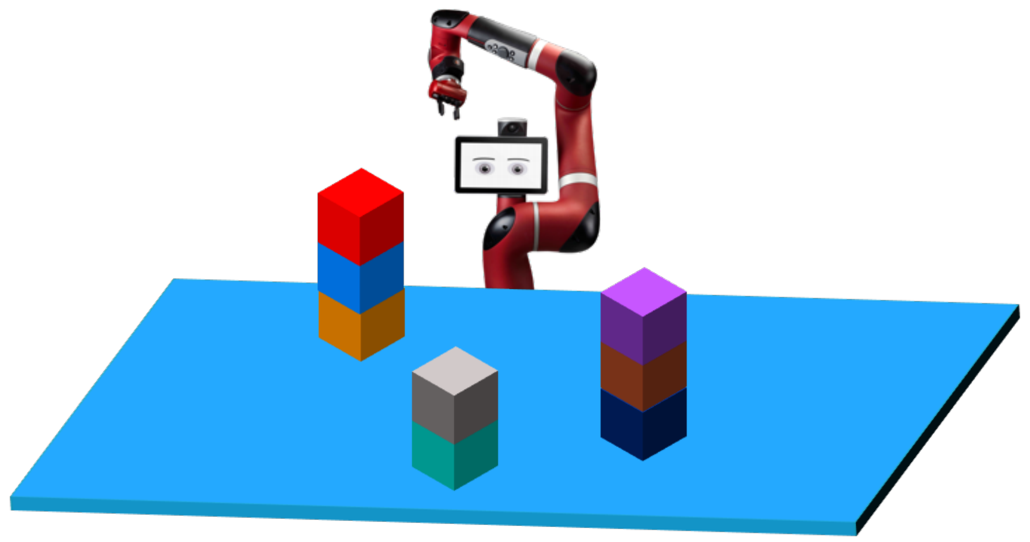
[Xu*, Nair*, et al. ICRA 2018; Huang*, Nair*, Xu*, et al. CVPR'2019; Huang et al. IROS'19]

One-Shot Imitation Learning from Videos: **Neural Task Graphs (NTG)**

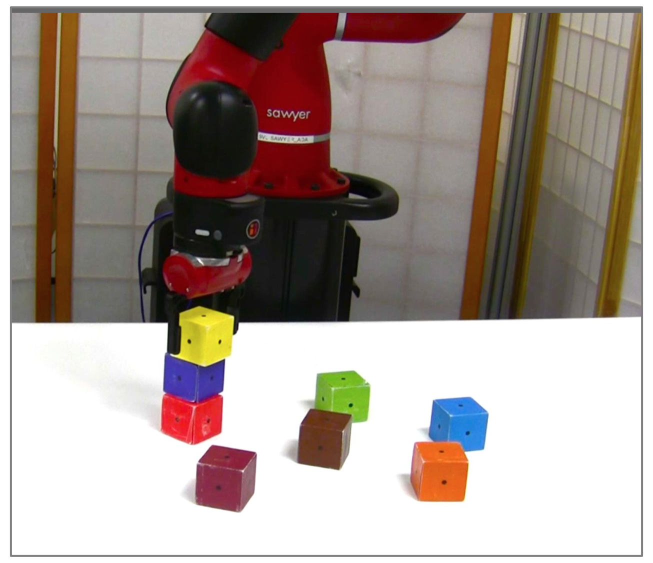


video demonstration

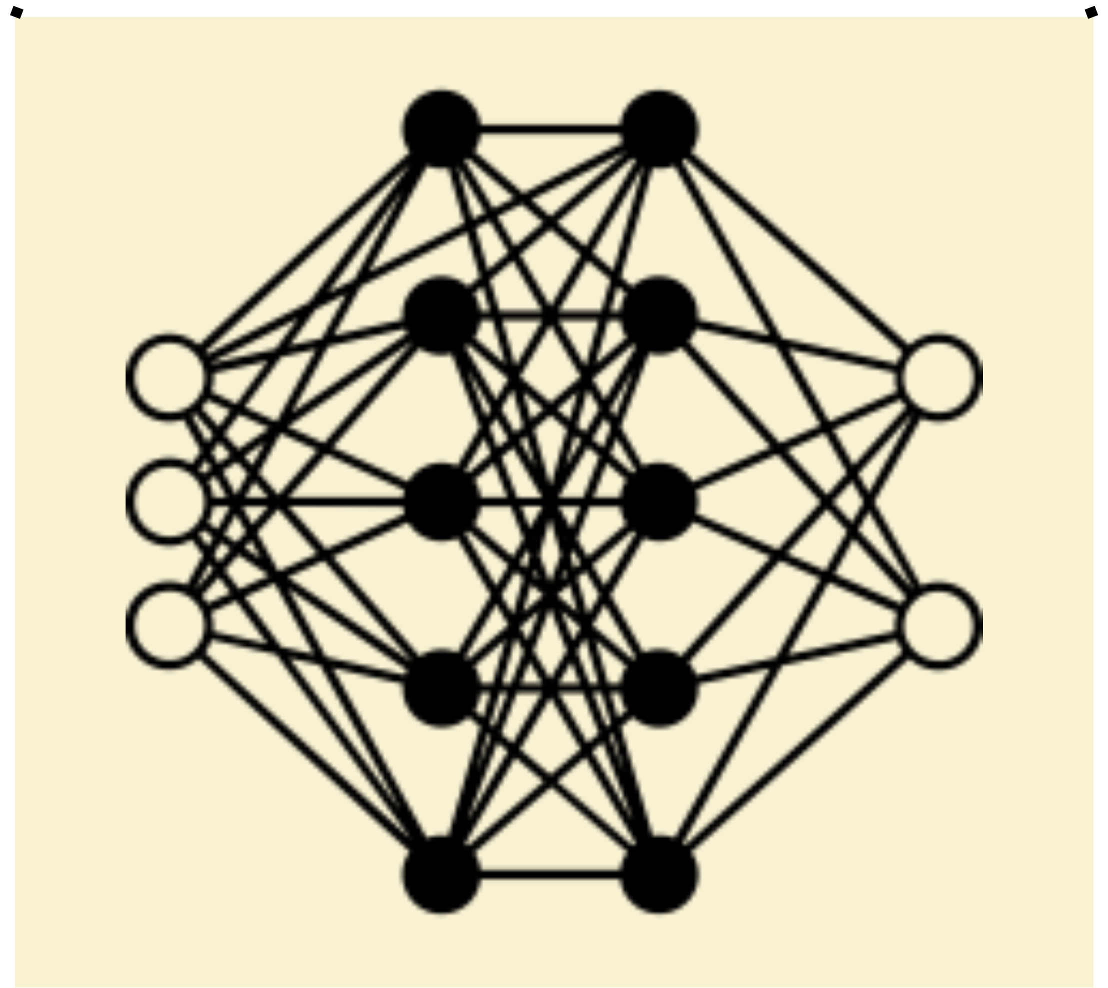
meta-policy



demo conditional policy



current observation

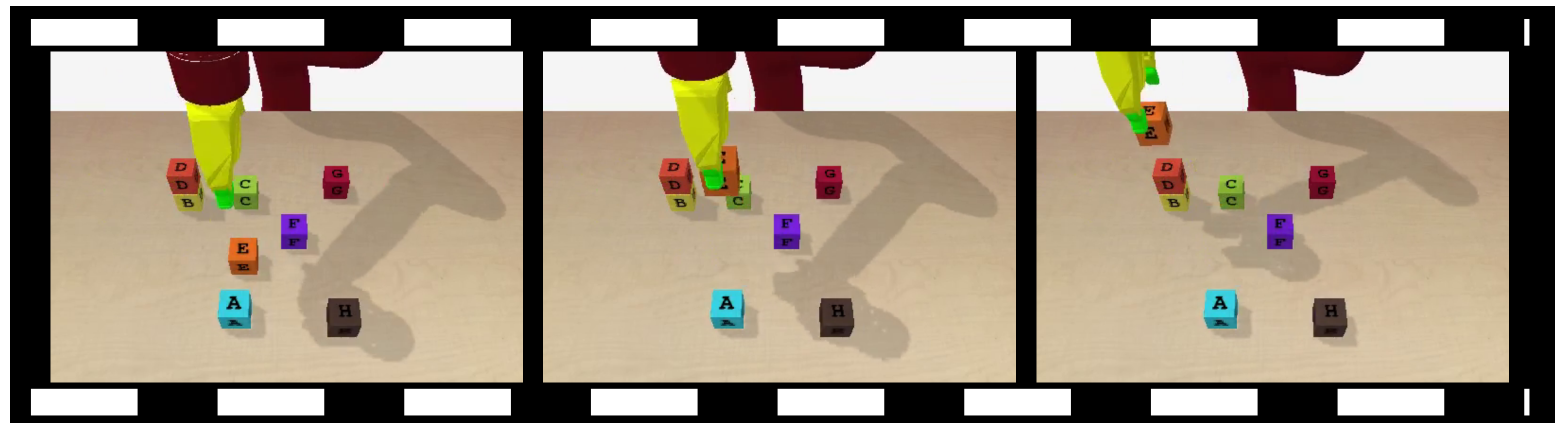


next action

pick (A)

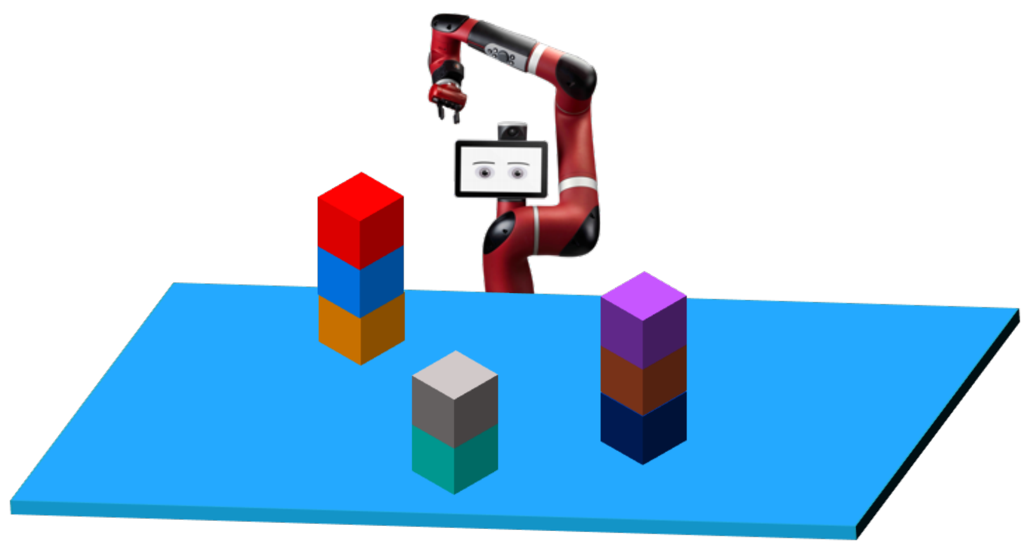
[Duan et al. 2017; Finn et al. 2017; Wang et al. 2017; Yu et al. 2018; Xu et al. 2018 "Neural Task Programming"]

One-Shot Imitation Learning from Videos: **Neural Task Graphs (NTG)**

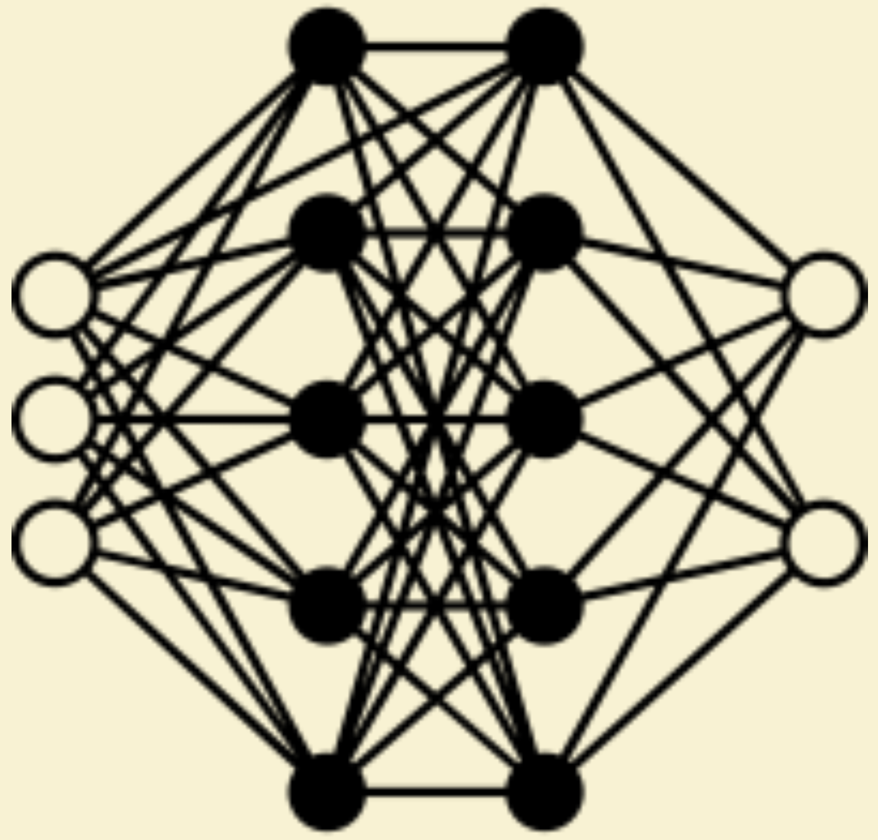


video demonstration

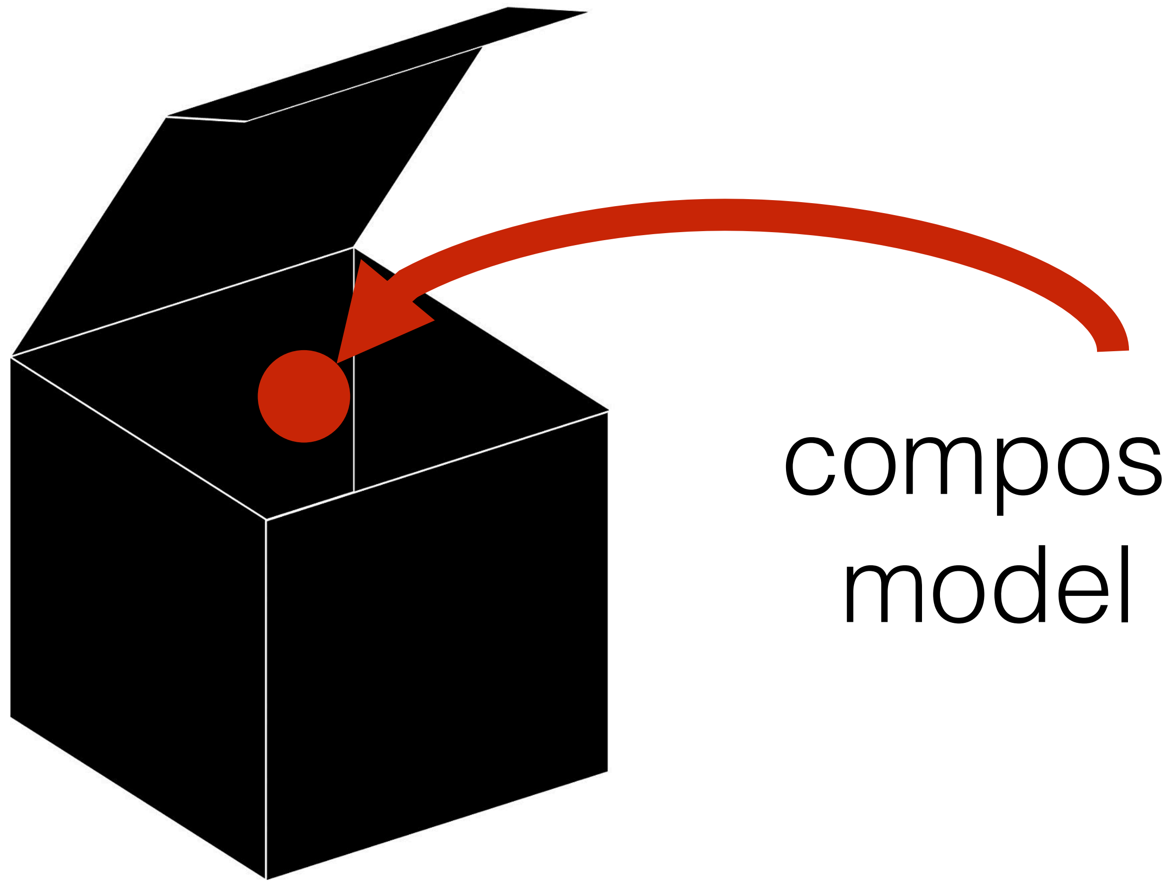
meta-policy



demo conditional policy



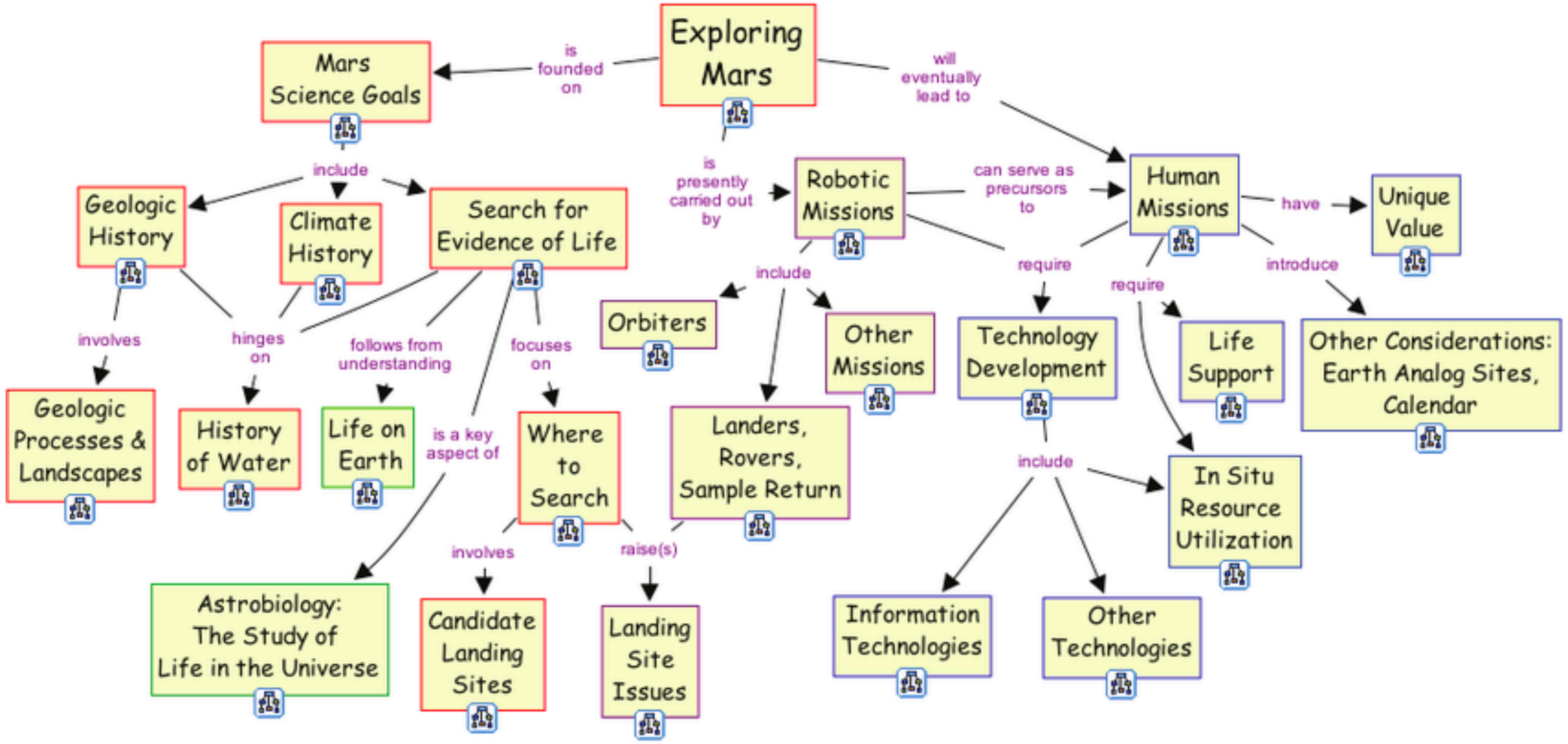
“the black box”



compositional
model prior

Key idea: Opening the Black Box

One-Shot Imitation Learning from Videos: **Neural Task Graphs (NTG)**



Task Graph created by NASA for Mars Exploration

Capturing compositional structures

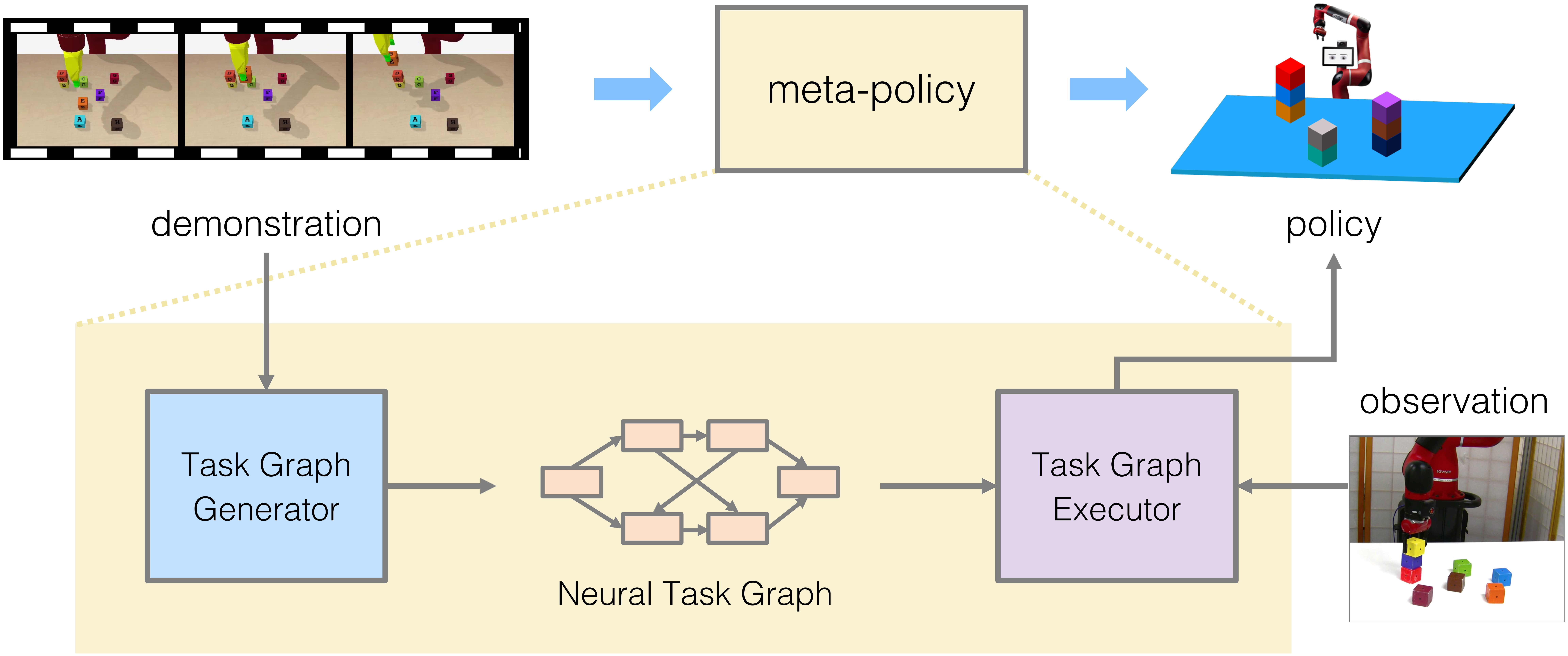
High degree of human interpretability

Questions

How to automatically construct these task graphs from video demonstrations?

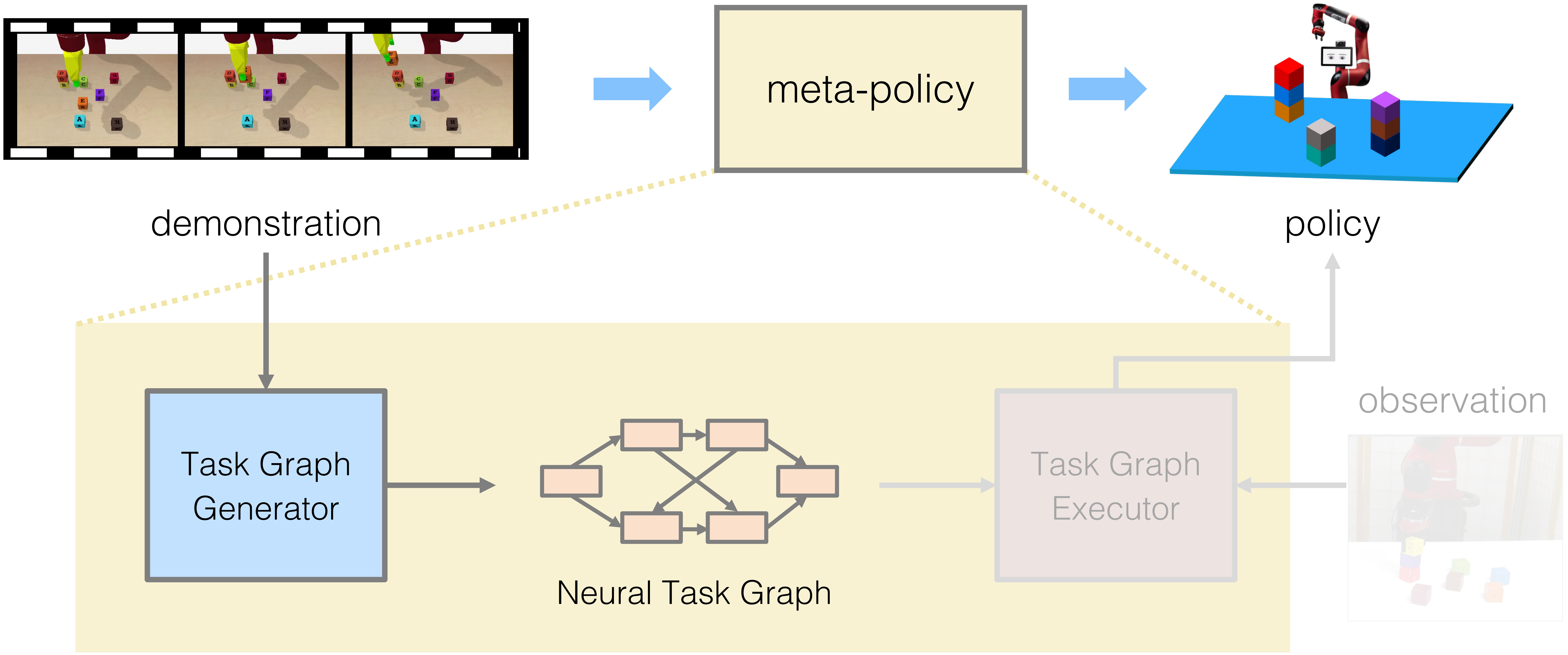
How to use these task graphs as model priors in deep learning methods?

One-Shot Imitation Learning from Videos: **Neural Task Graphs (NTG)**



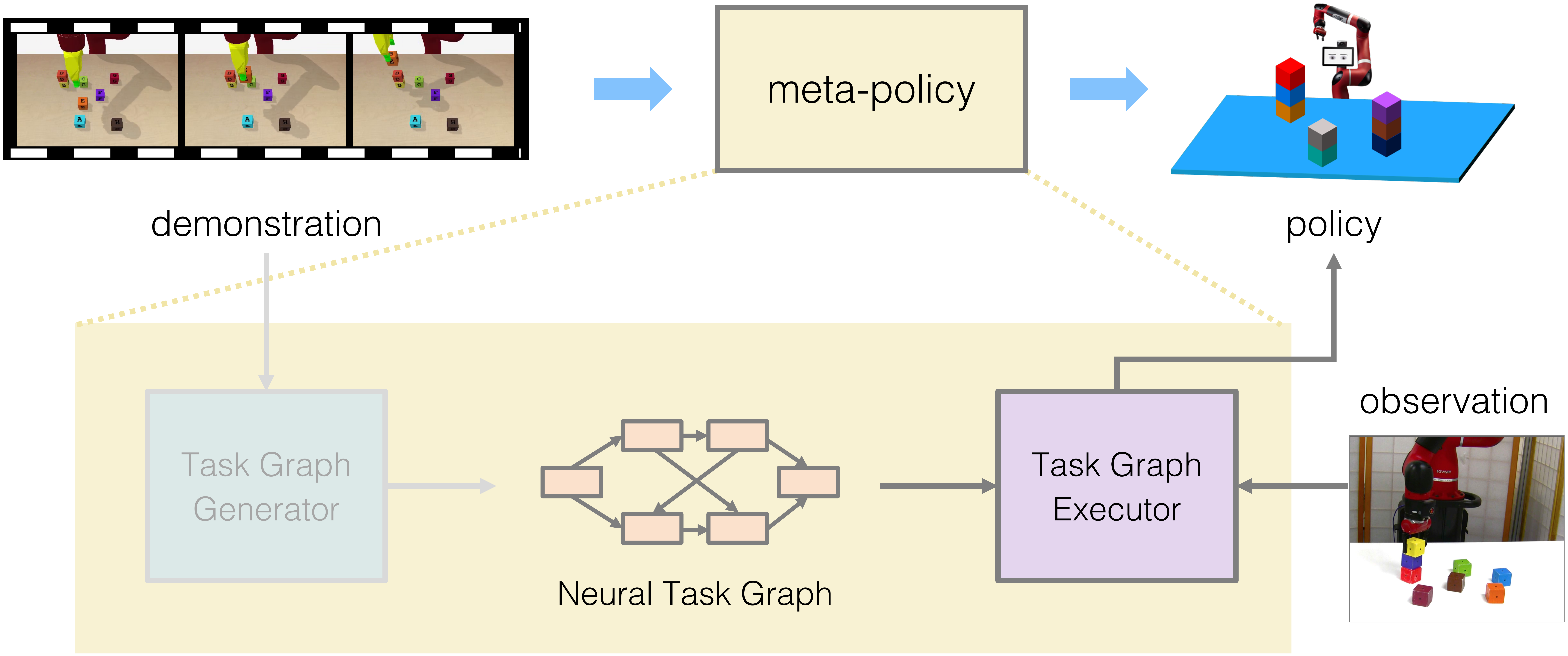
Huang*, Nair*, Xu*, et al. “Neural Task Graph”, CVPR 2019

One-Shot Imitation Learning from Videos: **Neural Task Graphs (NTG)**



Huang*, Nair*, Xu*, et al. “Neural Task Graph”, CVPR 2019

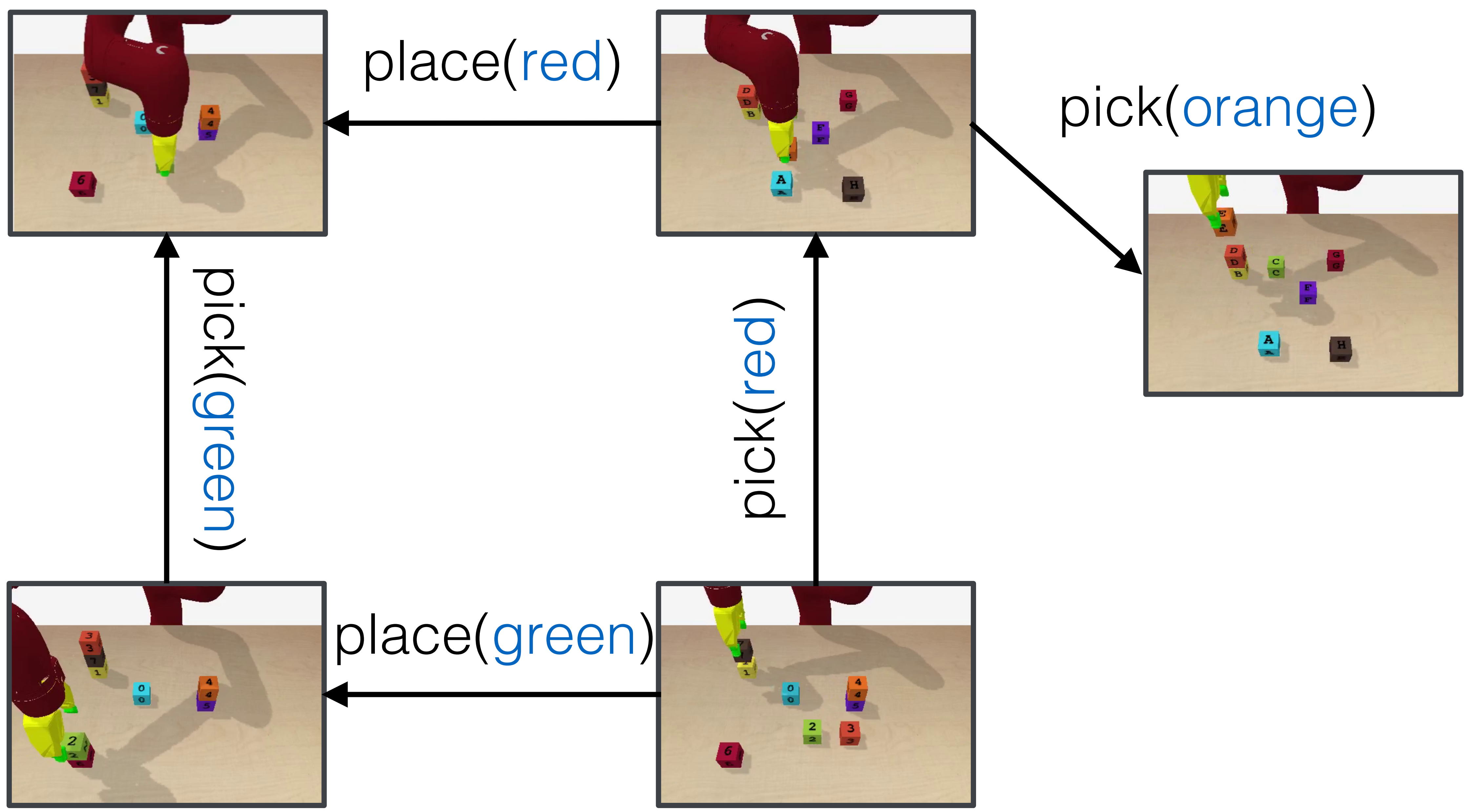
One-Shot Imitation Learning from Videos: **Neural Task Graphs (NTG)**



Huang*, Nair*, Xu*, et al. “Neural Task Graph”, CVPR 2019

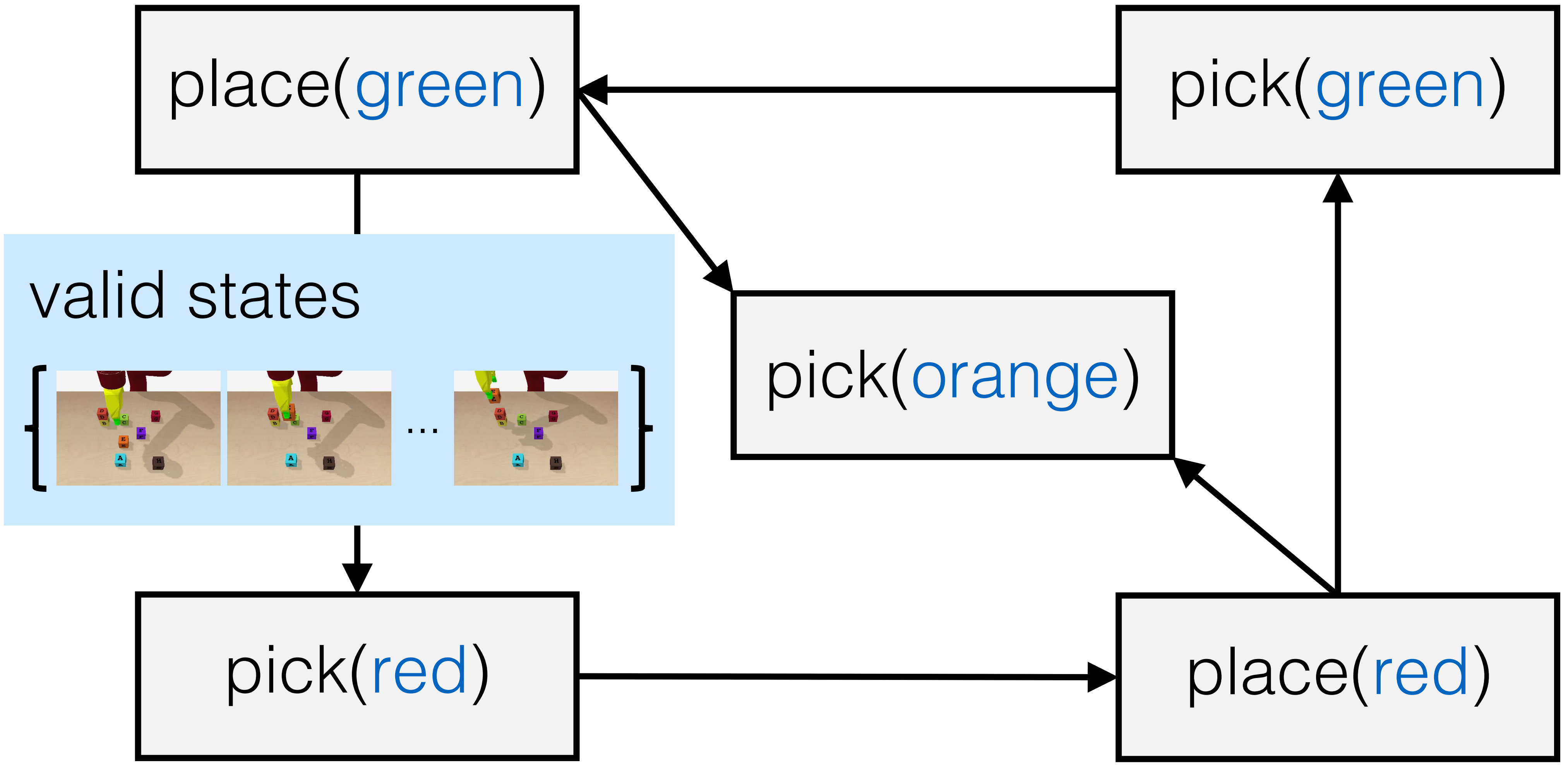
One-Shot Imitation Learning from Videos: **Neural Task Graphs (NTG)**

Task Graph



Nodes	States	combinatorial
Edges	Actions	

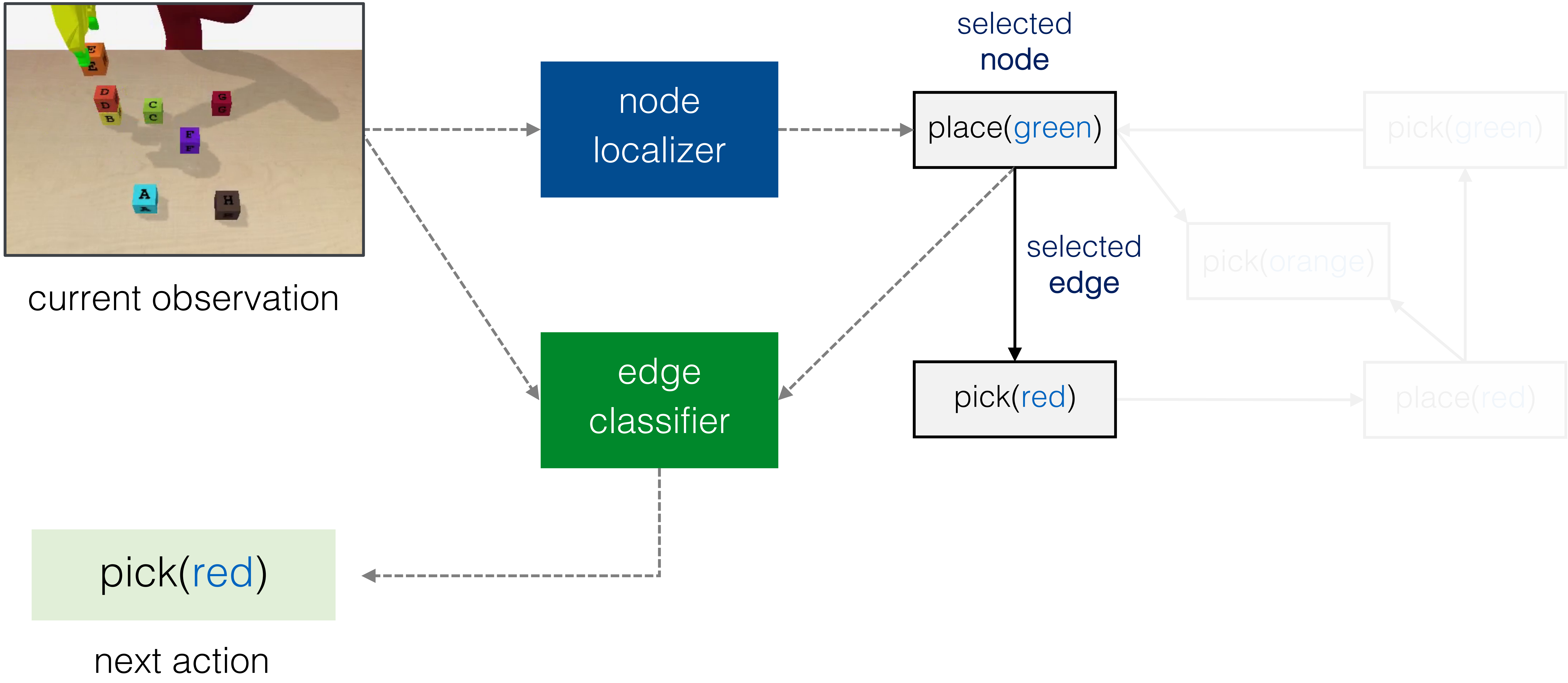
Conjugate Task Graph



Nodes	Actions	finite
Edges	States (Preconditions)	

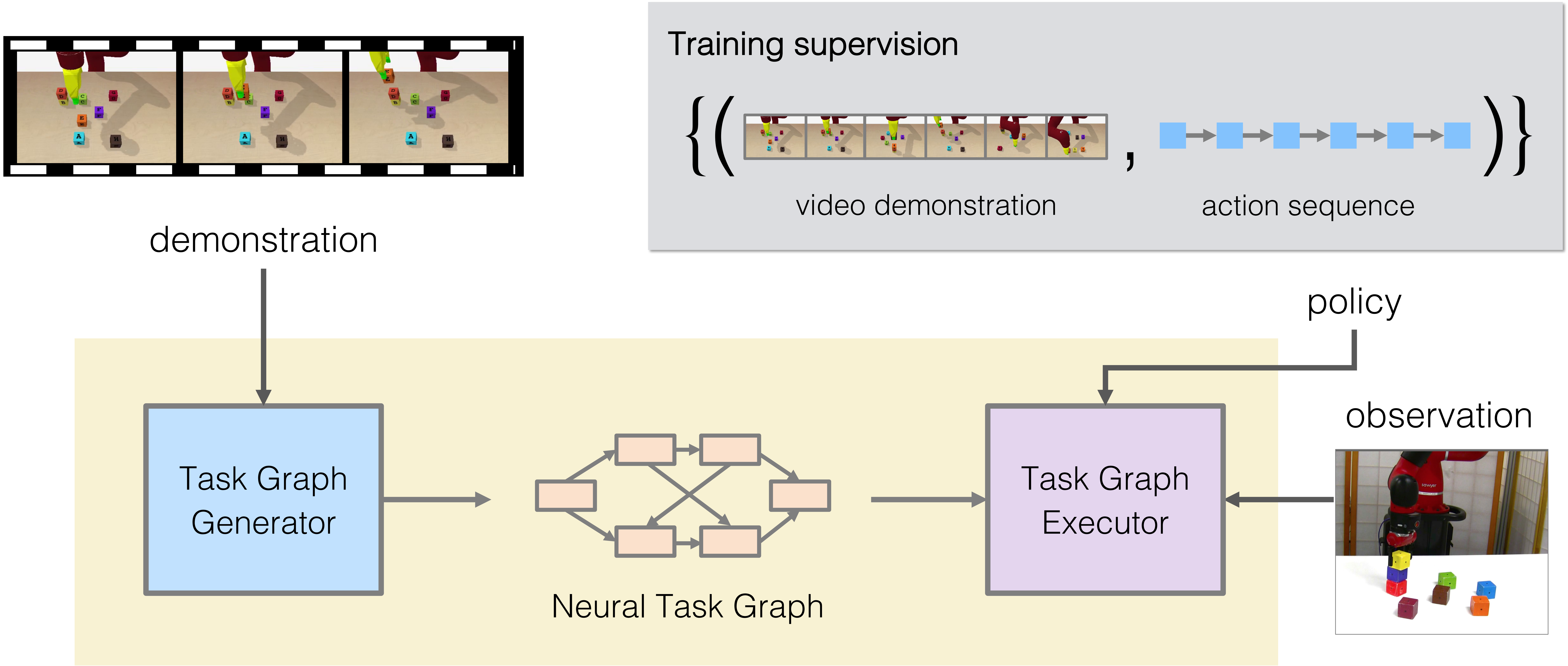
Huang*, Nair*, Xu*, et al. “Neural Task Graph”, CVPR 2019

One-Shot Imitation Learning from Videos: **Neural Task Graphs (NTG)**



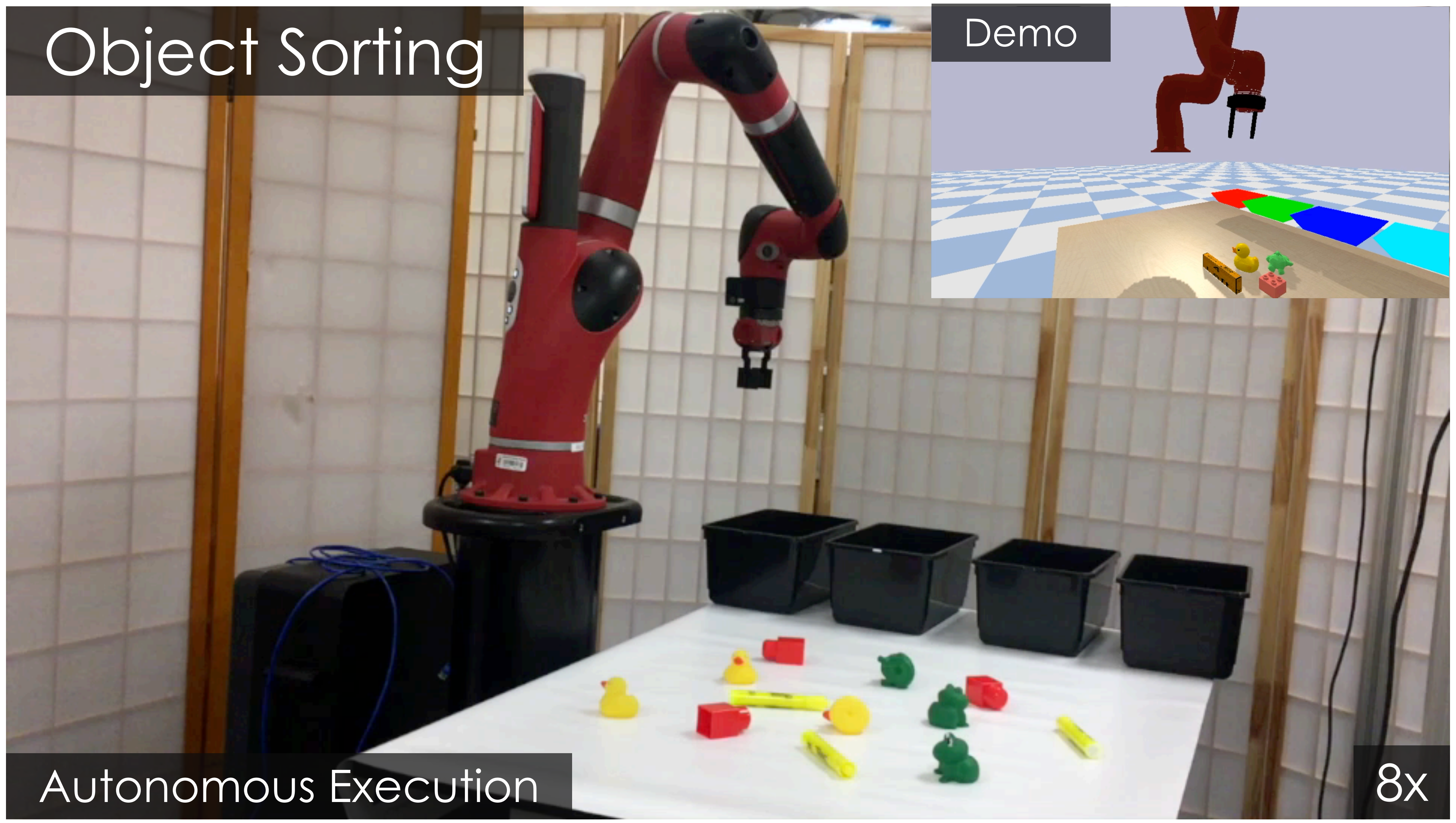
Huang*, Nair*, Xu*, et al. "Neural Task Graph", CVPR 2019

One-Shot Imitation Learning from Videos: **Neural Task Graphs (NTG)**

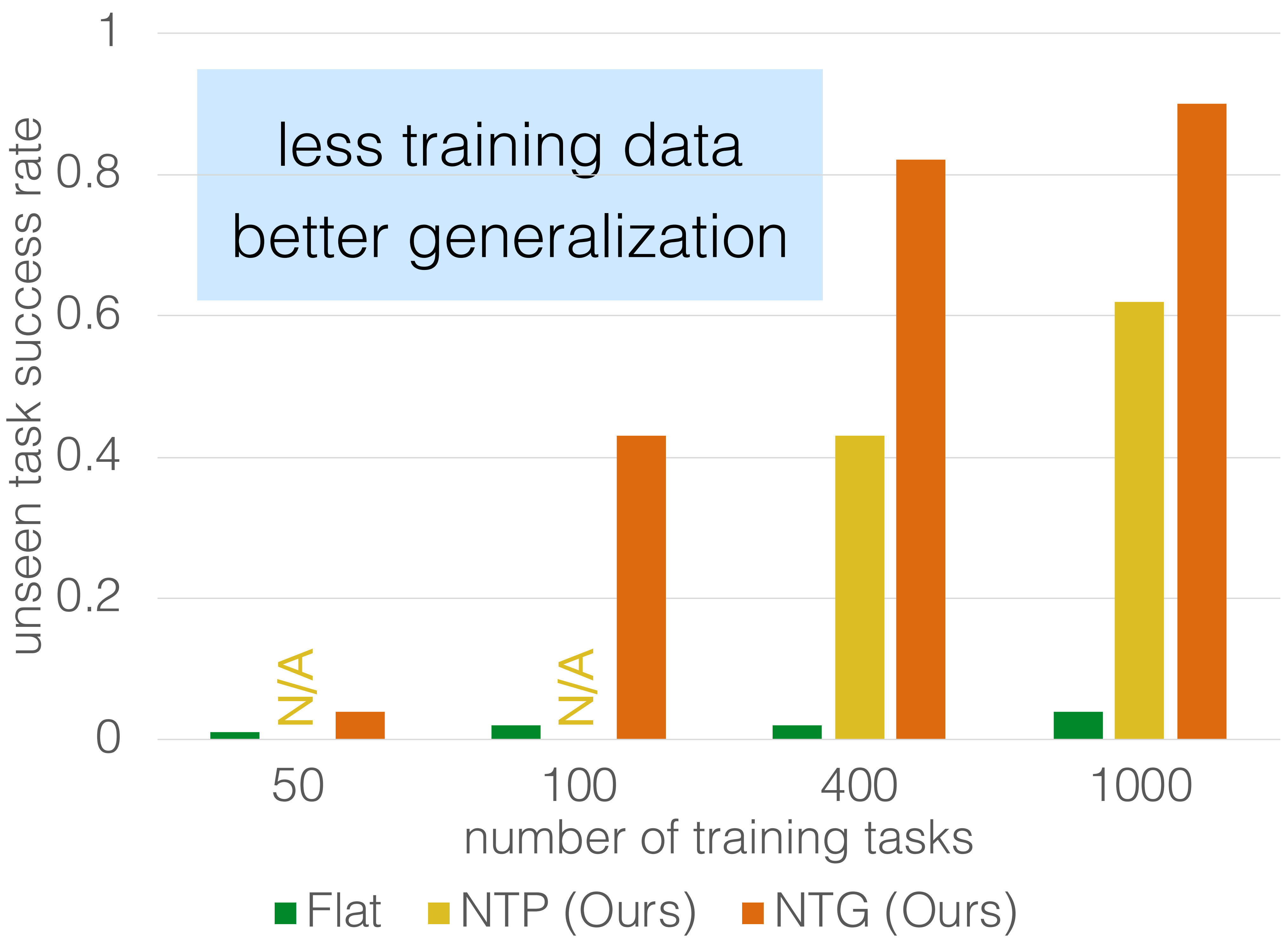


Huang*, Nair*, Xu*, et al. “Neural Task Graph”, CVPR 2019

One-Shot Imitation Learning from Videos: **Neural Task Graphs (NTG)**



Qualitative

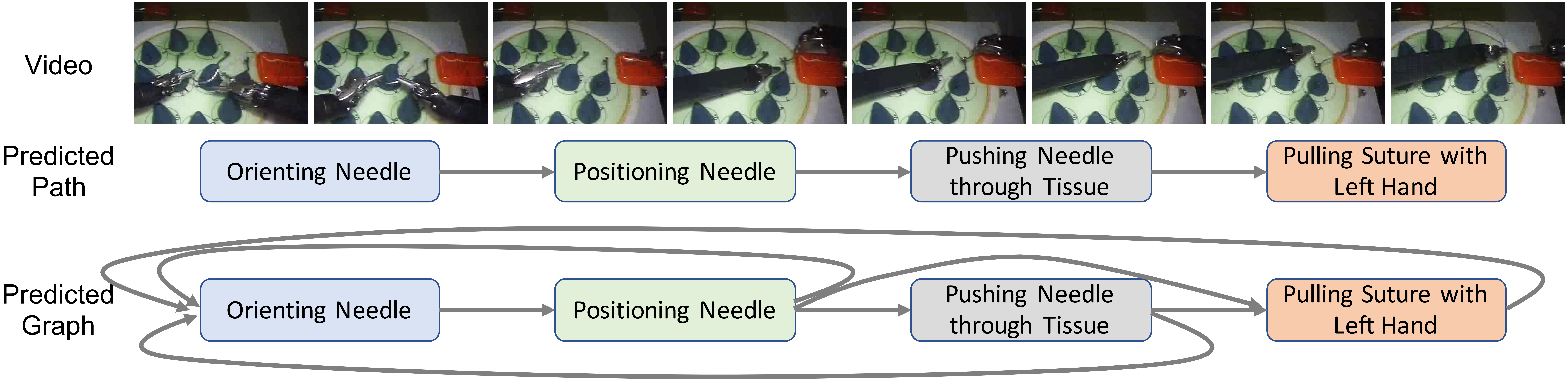


Quantitative

(the higher the better)

Huang*, Nair*, Xu*, et al. "Neural Task Graph", CVPR 2019

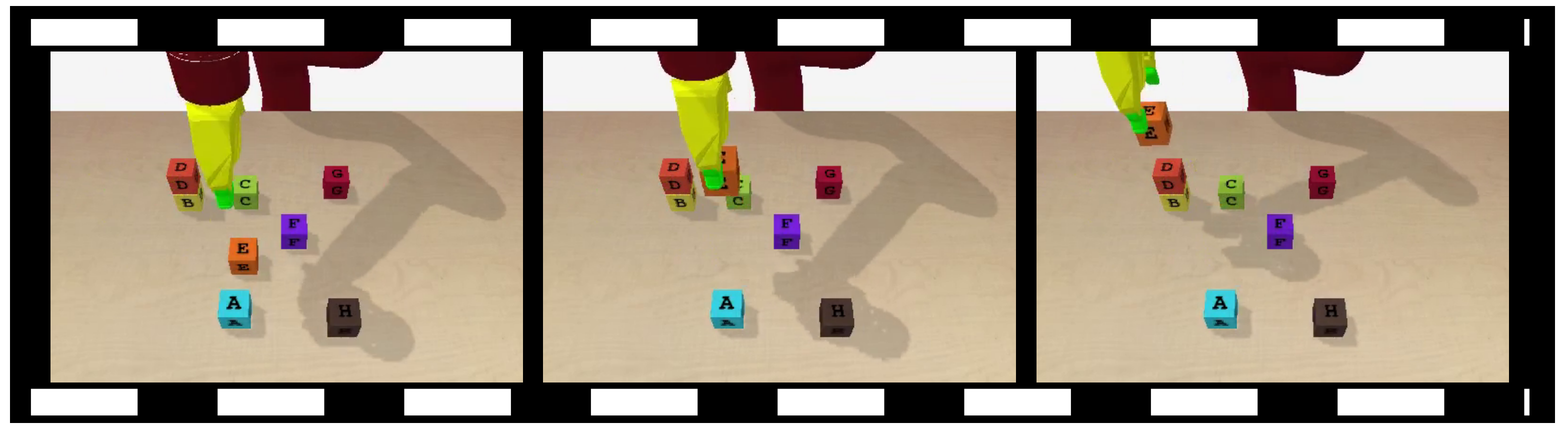
One-Shot Imitation Learning from Videos: **Neural Task Graphs (NTG)**



Applying NTG to the real-world surgical video dataset JIGSAWS

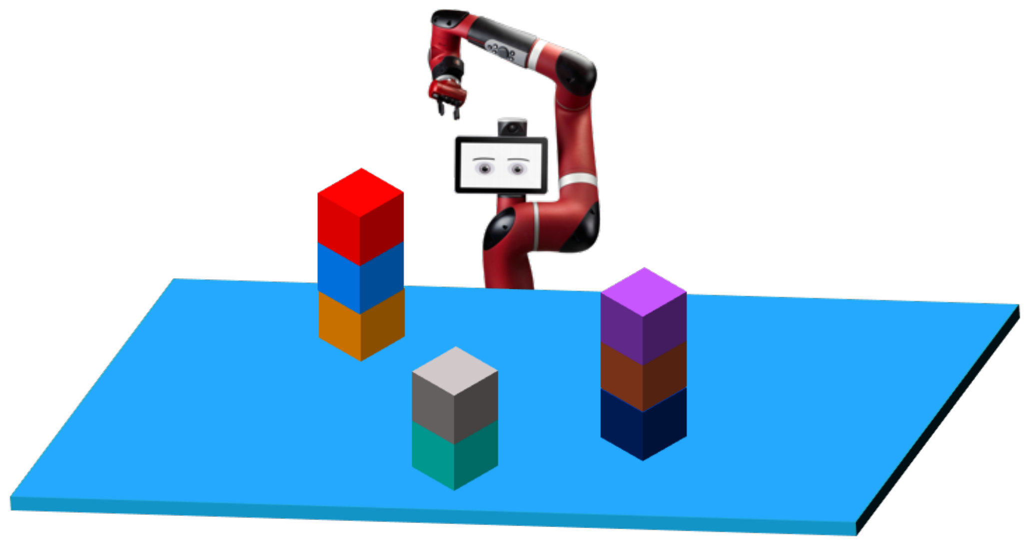
Huang*, Nair*, Xu*, et al. “Neural Task Graph”, CVPR 2019

One-Shot Imitation Learning from Videos: **Neural Task Graphs (NTG)**

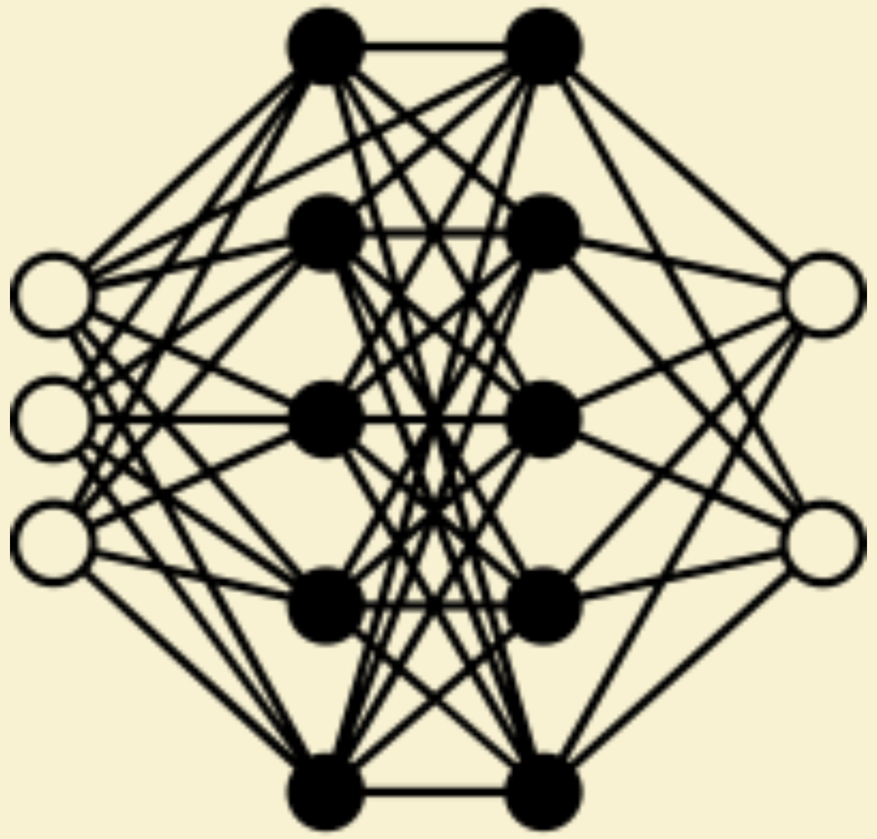


video demonstration

meta-policy

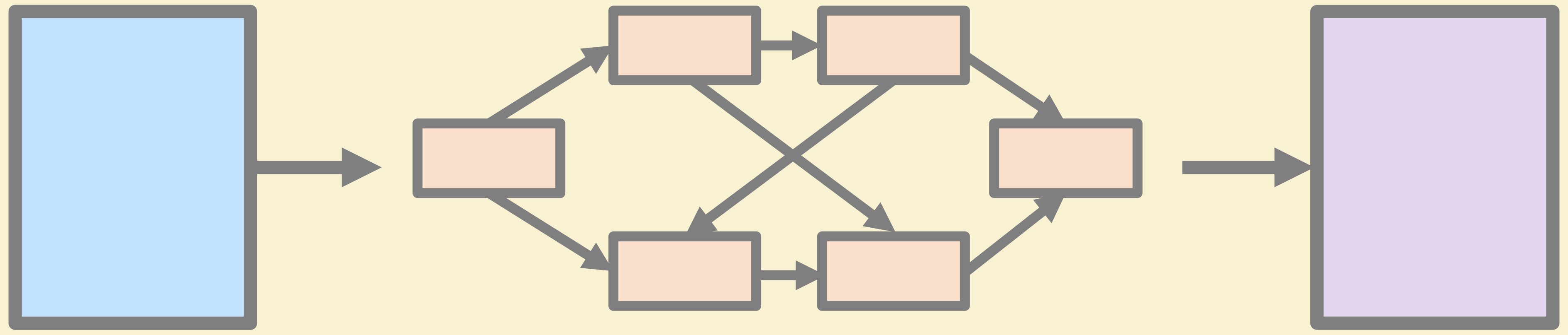


demo conditional policy



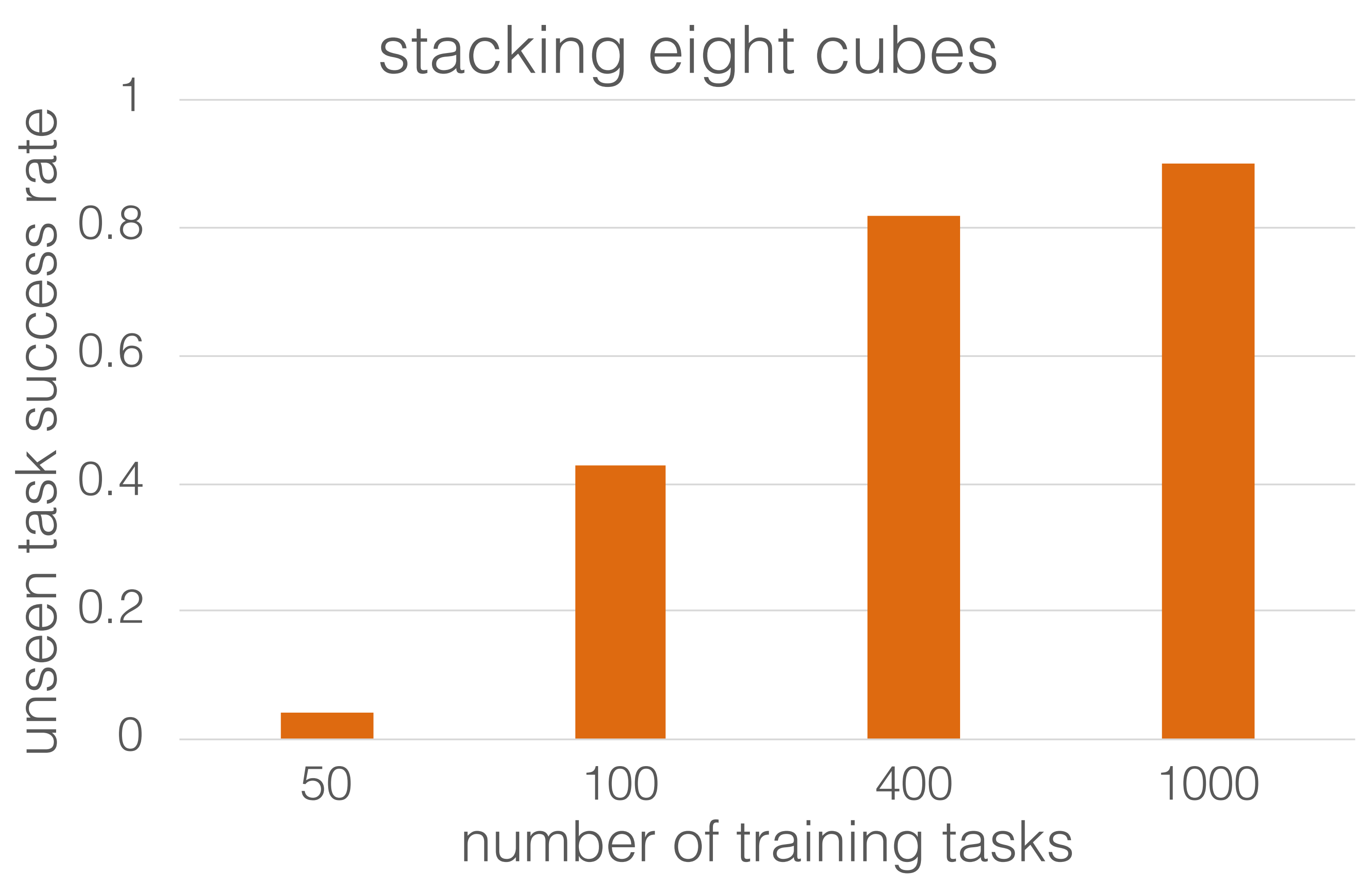
“the black box”

infuse
structural prior

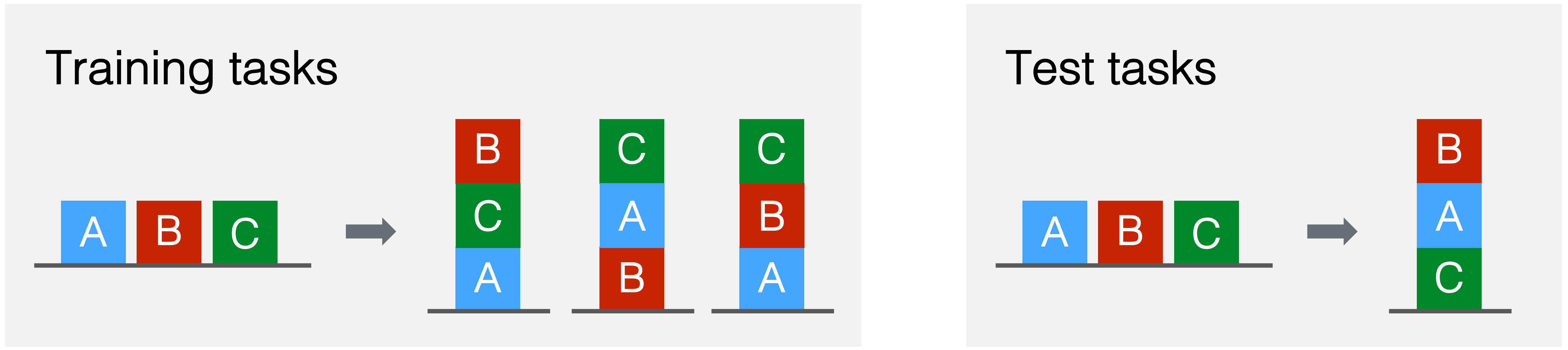


“model with compositional inductive bias”

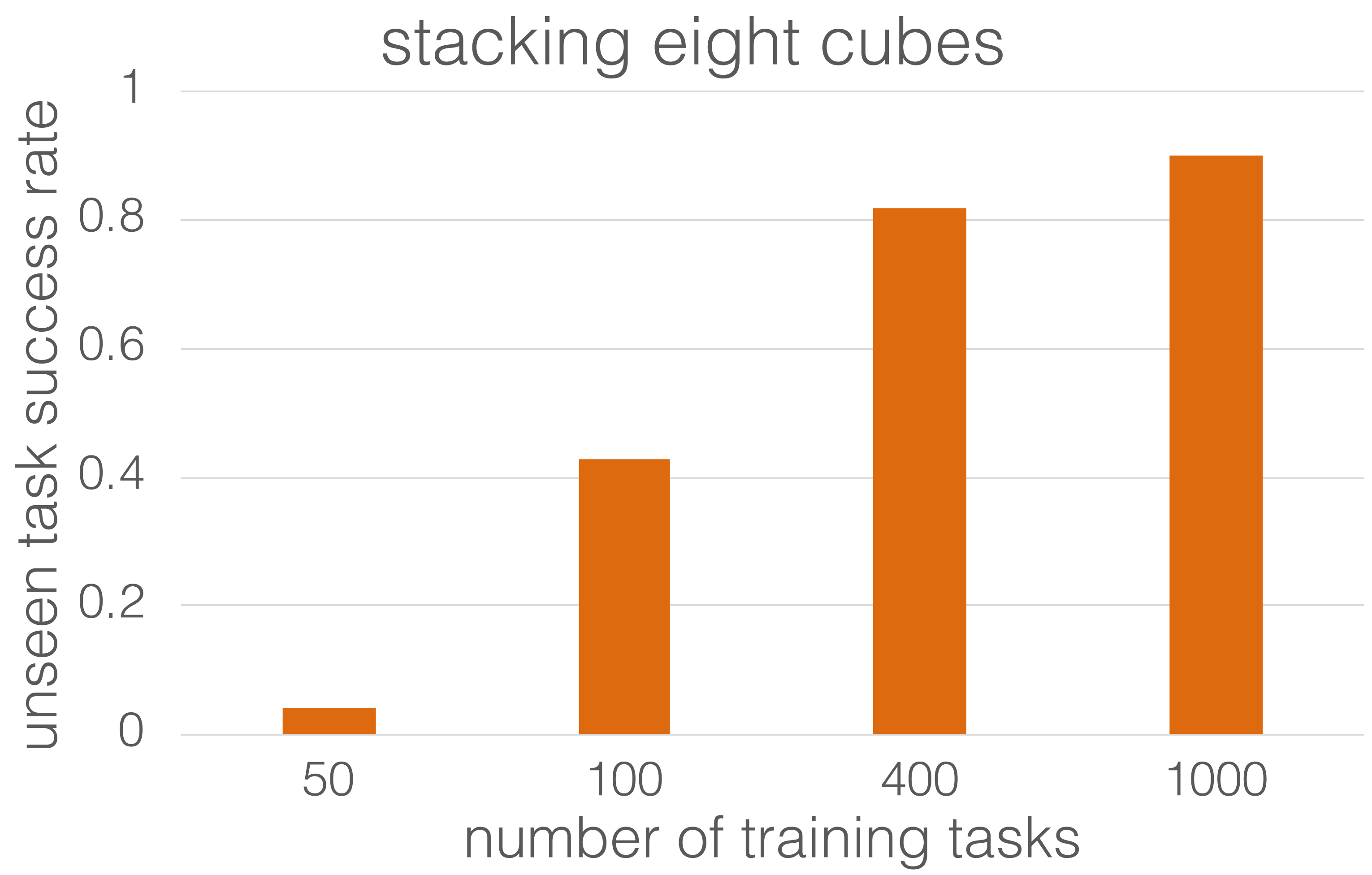
One-Shot Imitation Learning from Videos: **Neural Task Graphs (NTG)**



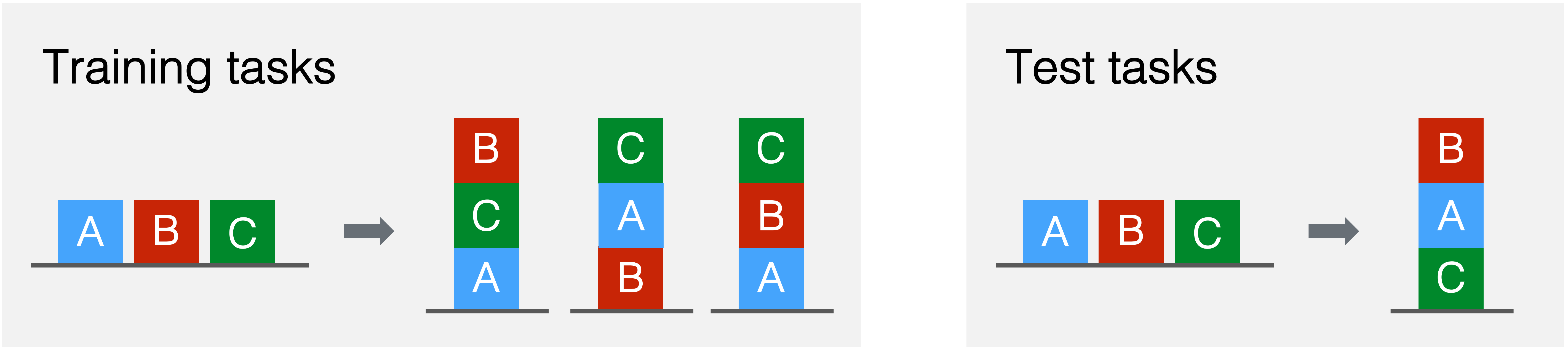
Compositional inductive bias in deep models is very effective for "in-distribution" task generalization (interpolation).



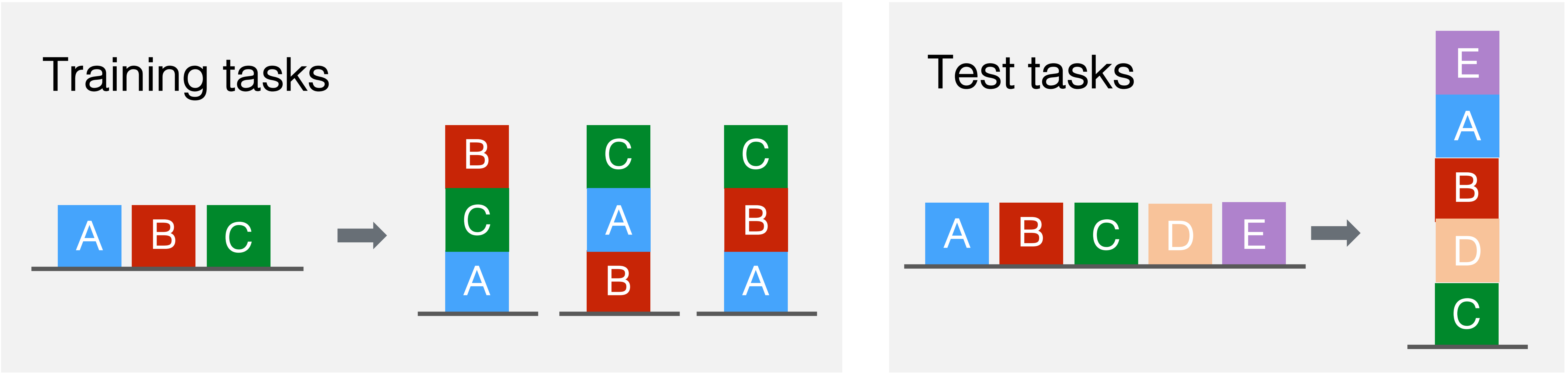
One-Shot Imitation Learning from Videos: **Neural Task Graphs (NTG)**

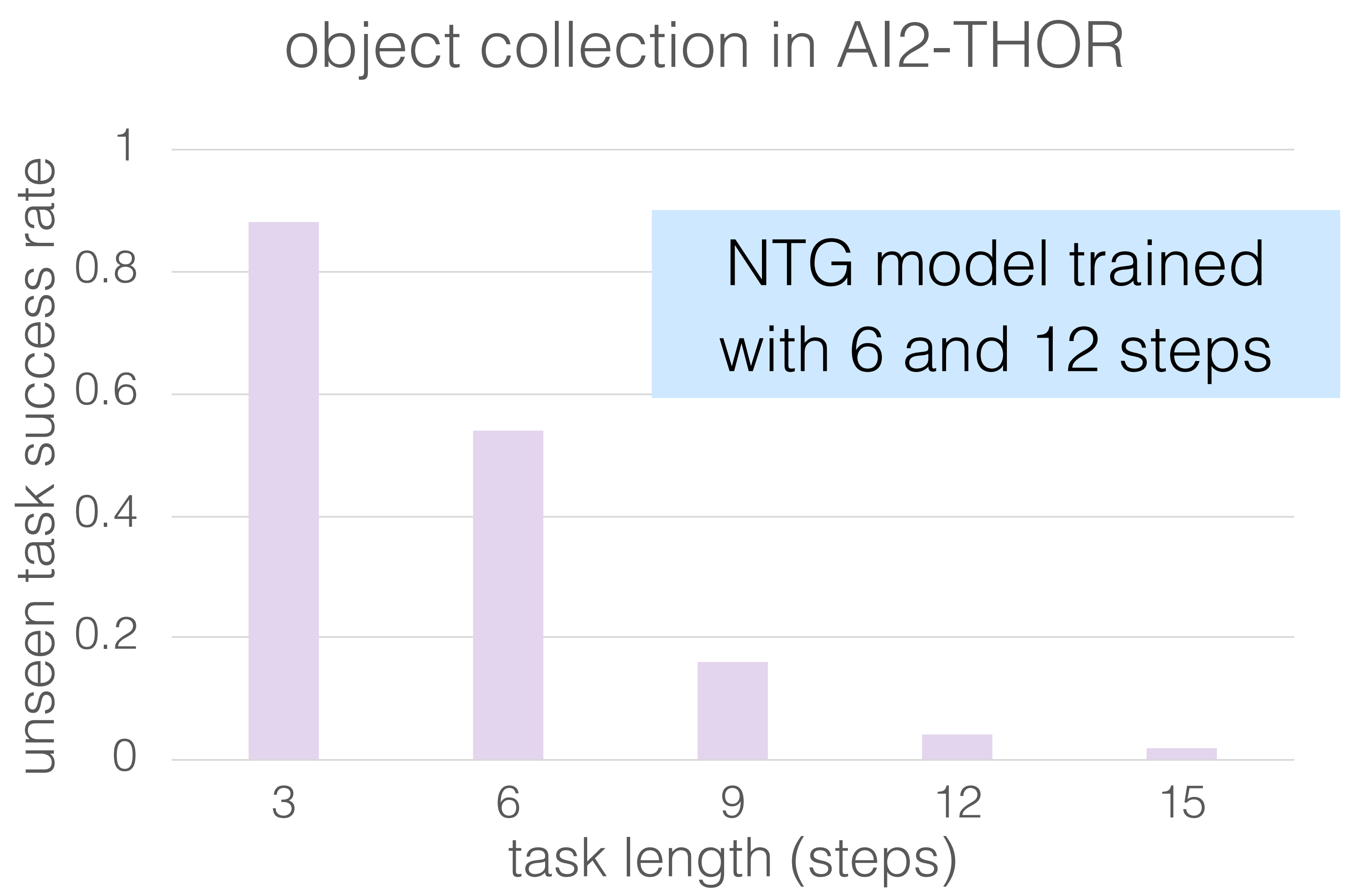


Compositional inductive bias in deep models is very effective for "in-distribution" task generalization (interpolation).



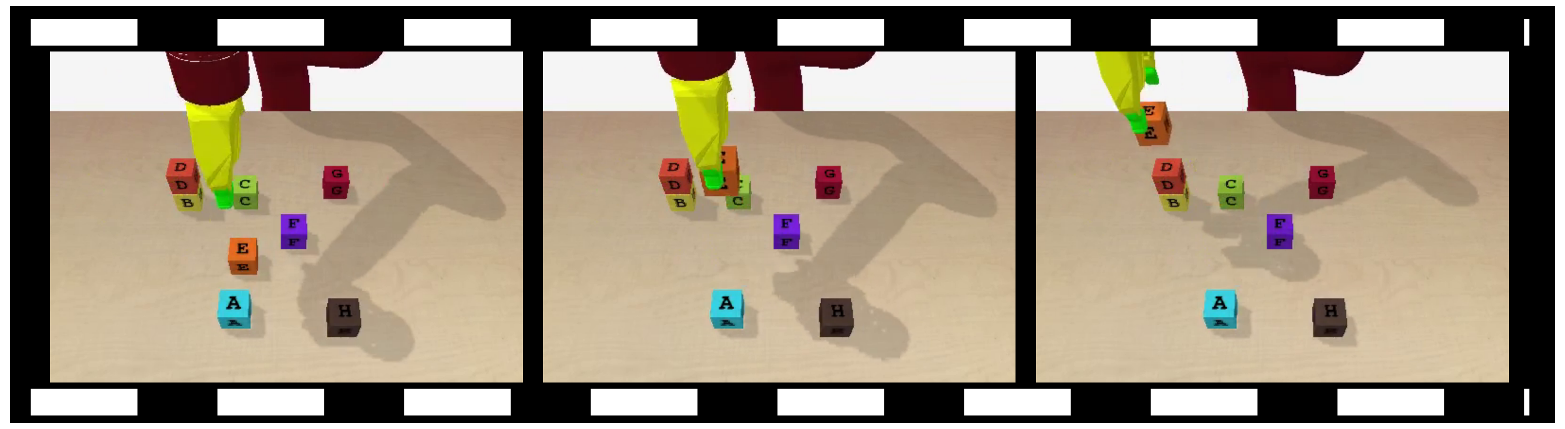
NTG is data-hungry and insufficient for "out-of-distribution" task generalization (extrapolation).





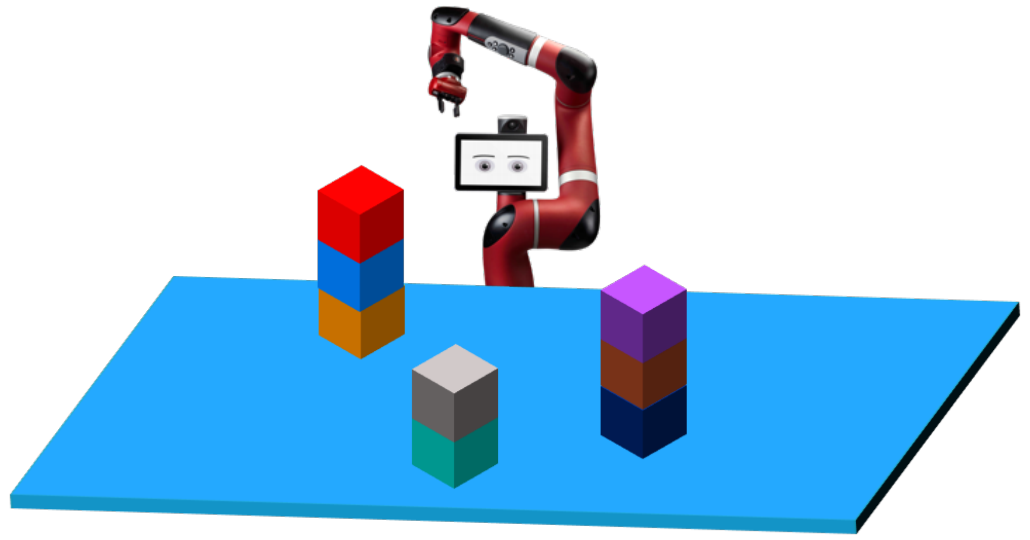
Visual Imitation Learning

One-shot imitation with **stronger task generalization**

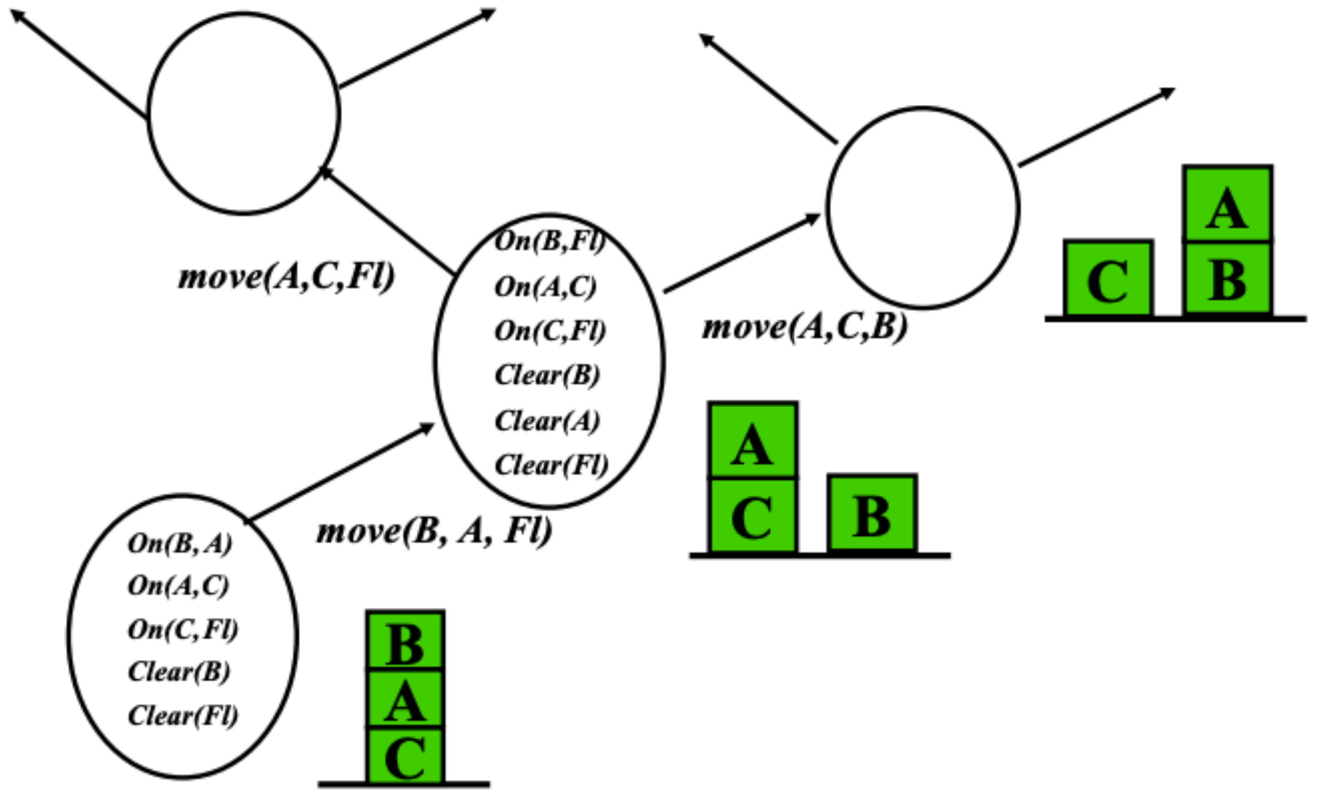


video demonstration

meta-policy



demo conditional policy



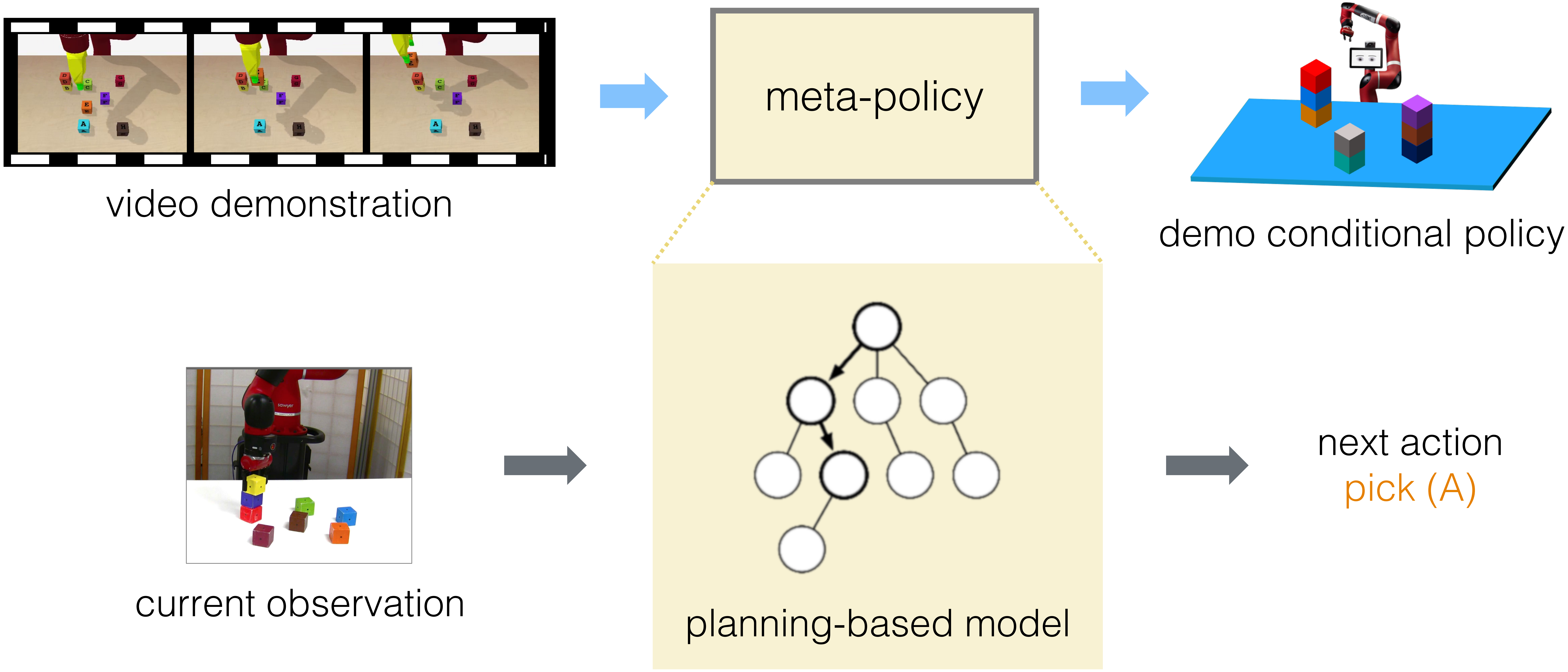
Classic symbolic planning (with additional domain knowledge) is capable of strong generalization.

Blocks Word Domain File

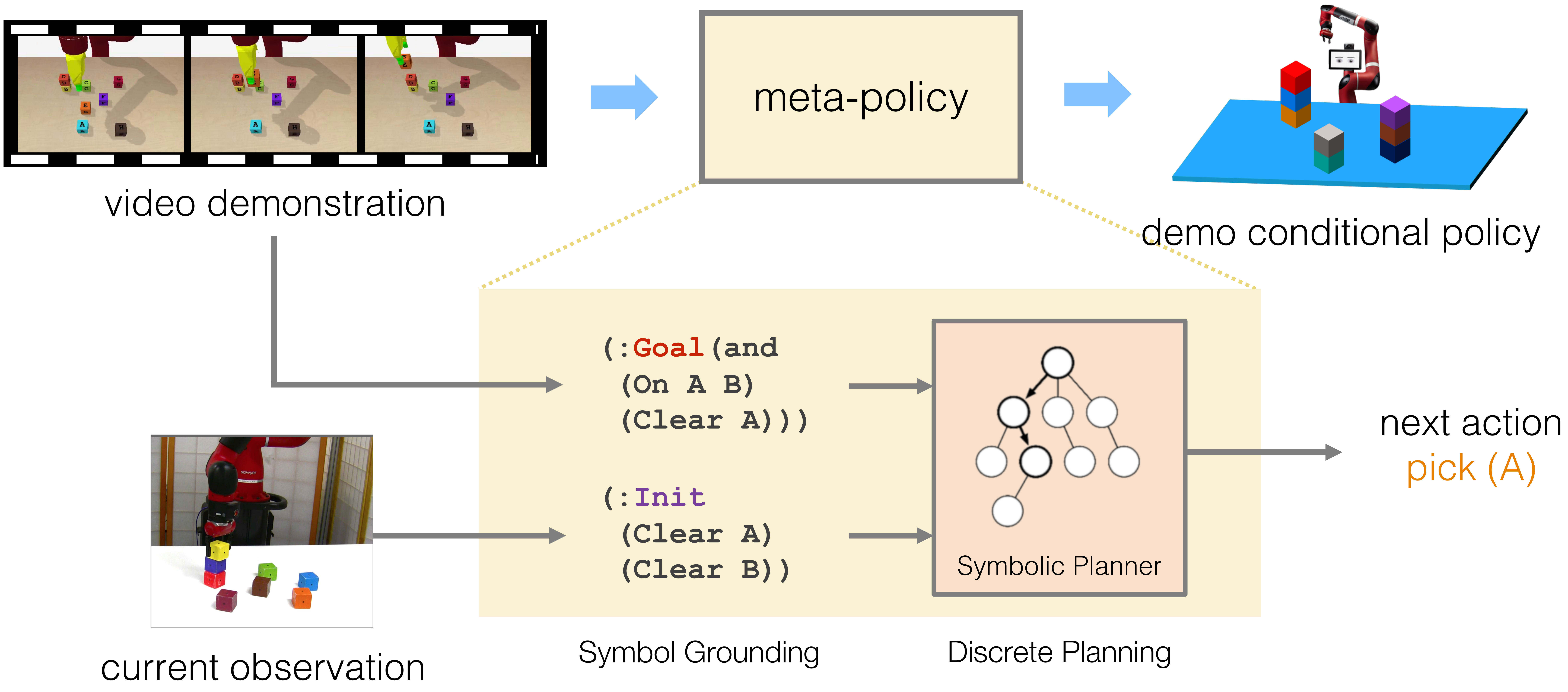
```
(define (domain hw5)
  (:requirements :strips)
  (:constants red green blue yellow)
  (:predicates (on ?x ?y) (on-table ?x) (block ?x) ... (clean ?x))
  (:action pick-up
    :parameters (?obj1)
    :precondition (and (clear ?obj1) (on-table ?obj1)
                       (arm-empty))
    :effect (and (not (on-table ?obj1))
                 (not (clear ?obj1))
                 (not (arm-empty))
                 (holding ?obj1)))
  ... more actions ...)
```

Visual Imitation Learning

One-shot imitation with **stronger task generalization**

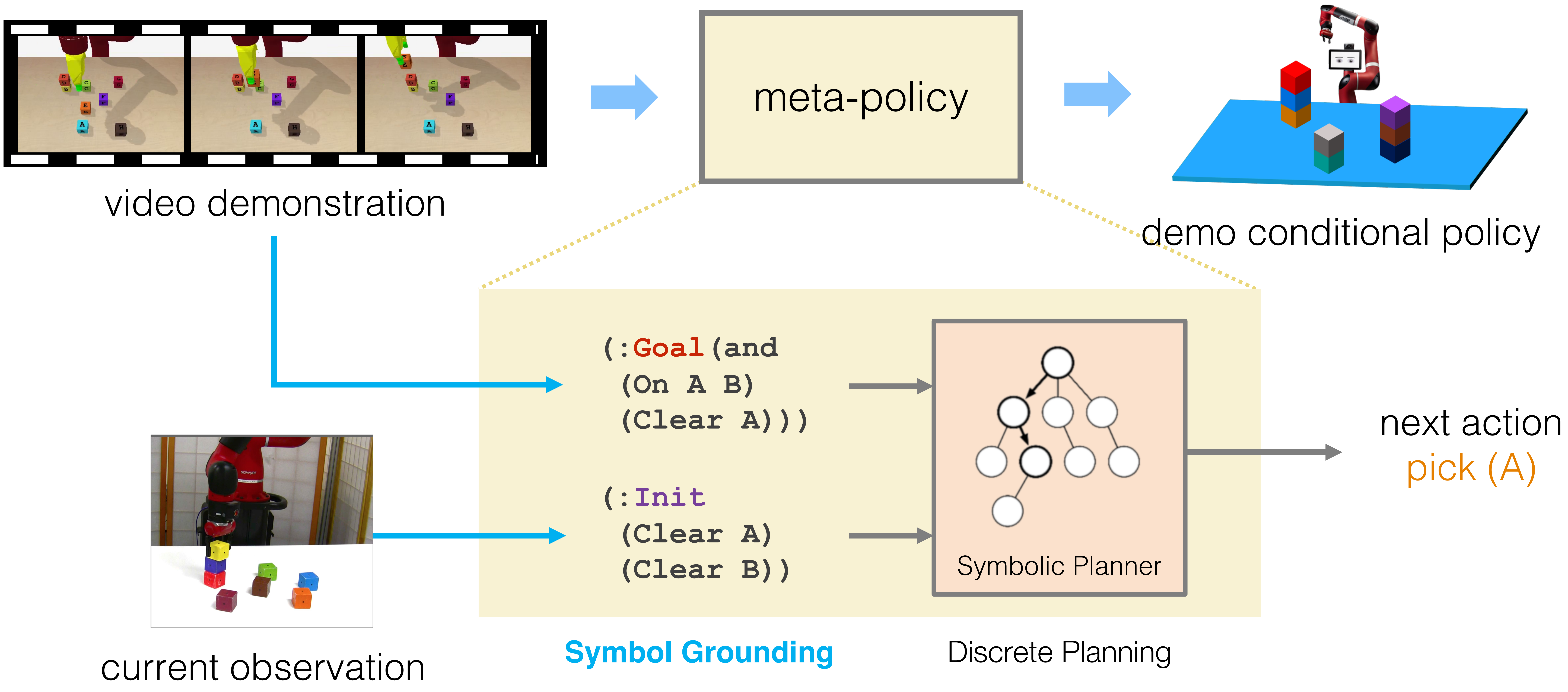


Planning-based One-Shot Visual Imitation: **Continuous Planner**



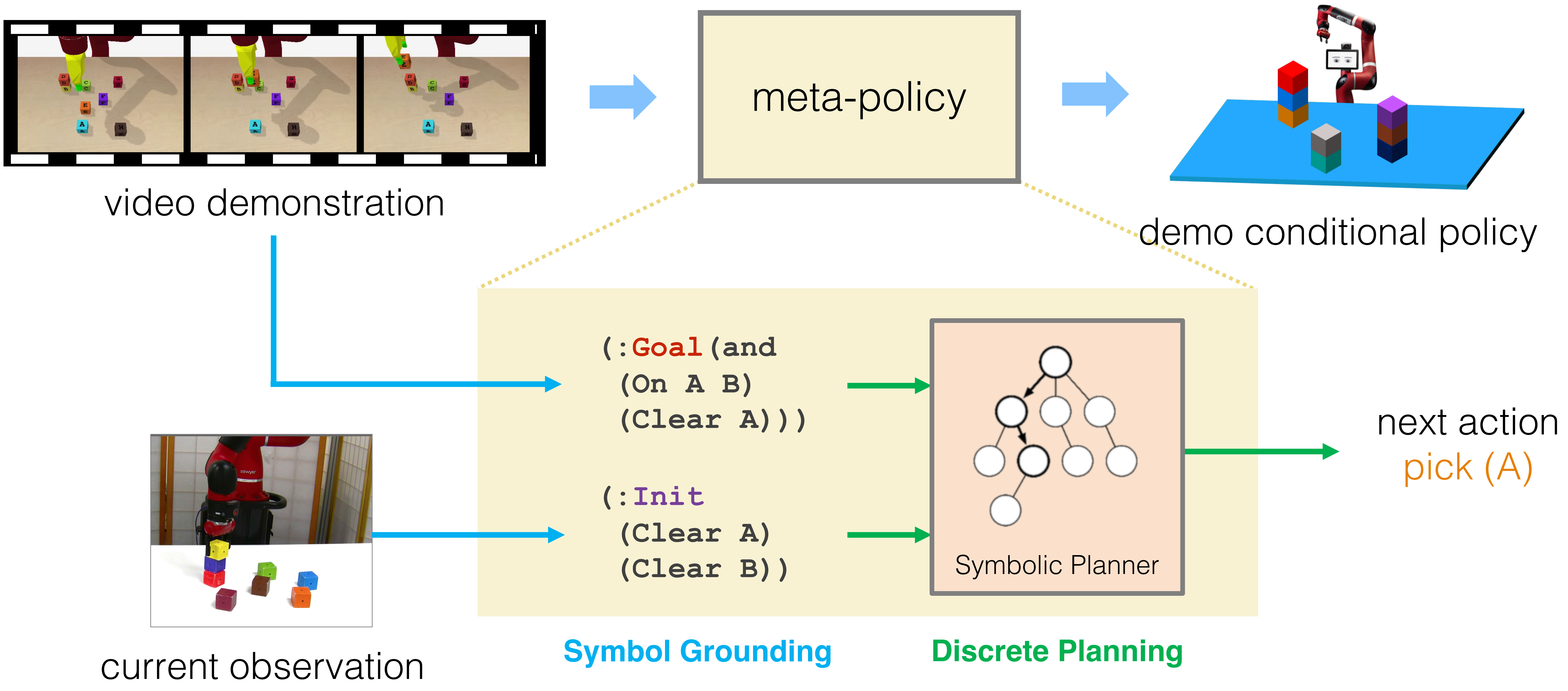
Huang et al. "Continuous Relaxation of Symbolic Planner", IROS 2019

Planning-based One-Shot Visual Imitation: **Continuous Planner**



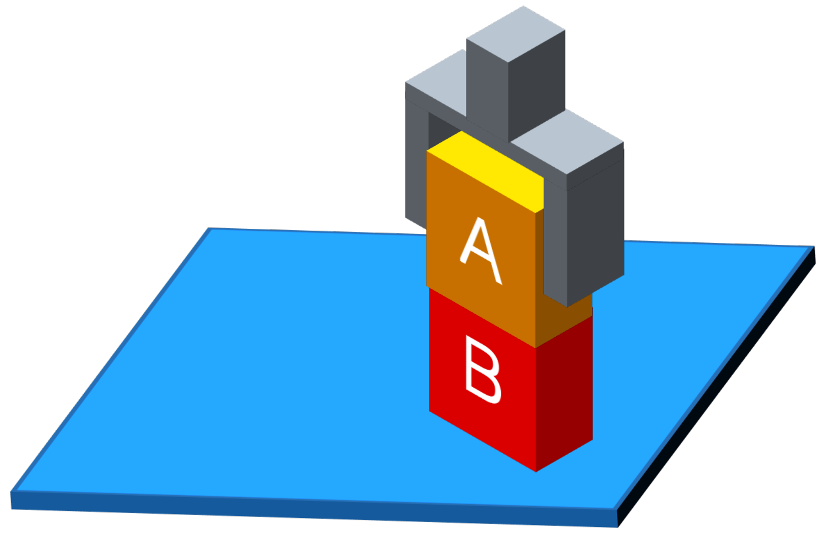
Huang et al. “Continuous Relaxation of Symbolic Planner”, IROS 2019

Planning-based One-Shot Visual Imitation: **Continuous Planner**



Huang et al. “Continuous Relaxation of Symbolic Planner”, IROS 2019

Planning-based One-Shot Visual Imitation: **Continuous Planner**



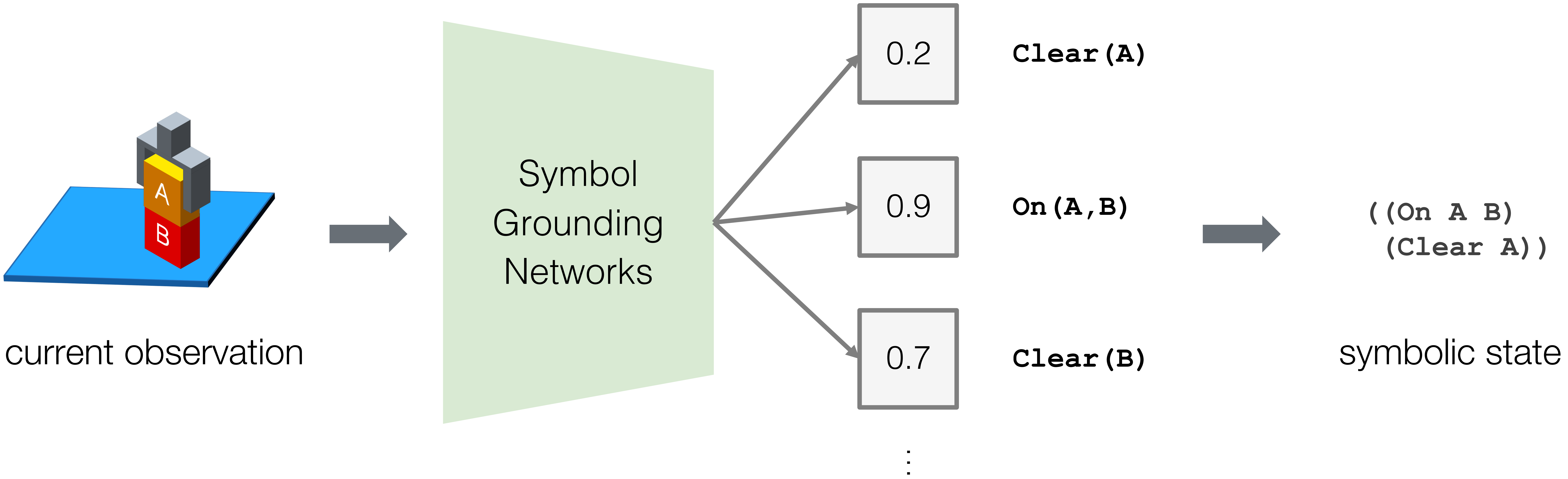
current observation

```
((On A B)
 (Clear A))
```

symbolic state

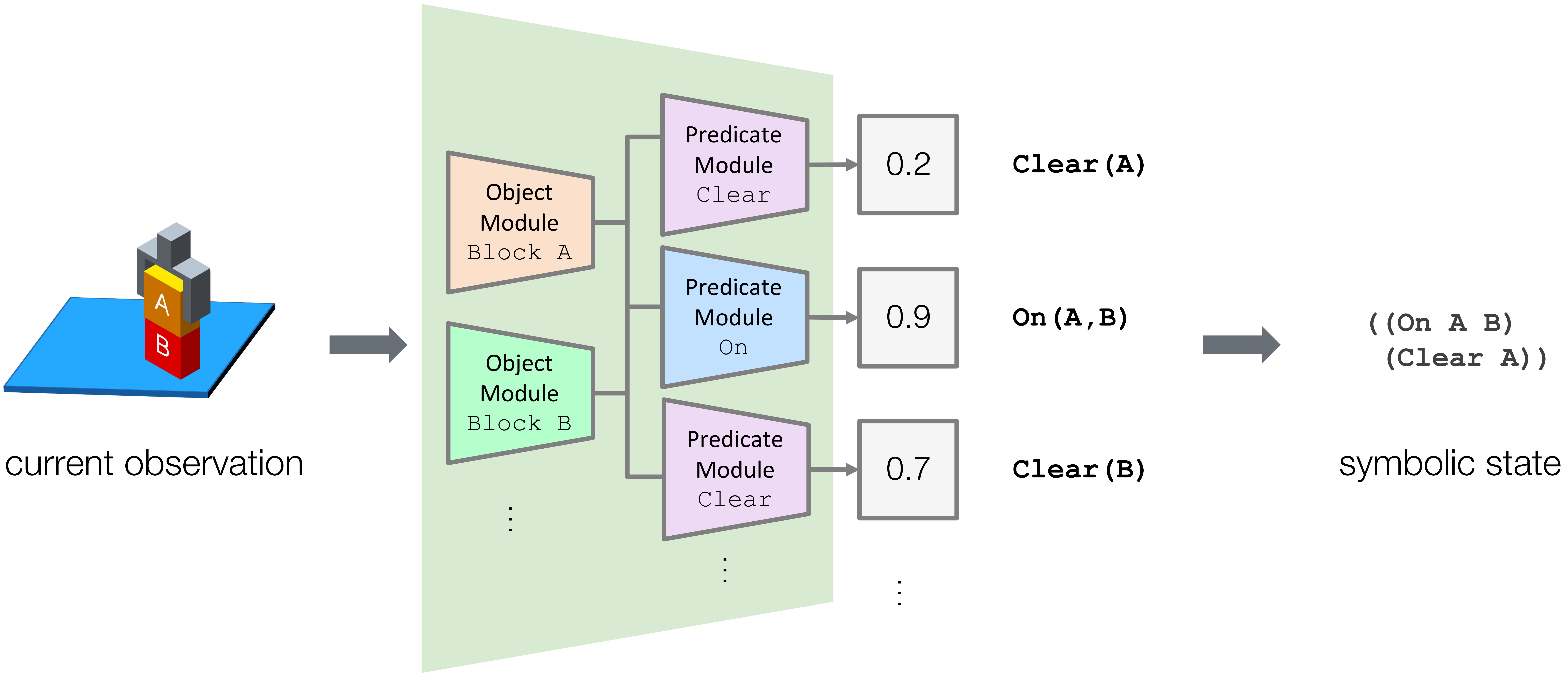
Huang et al. “Continuous Relaxation of Symbolic Planner”, IROS 2019

Planning-based One-Shot Visual Imitation: **Continuous Planner**



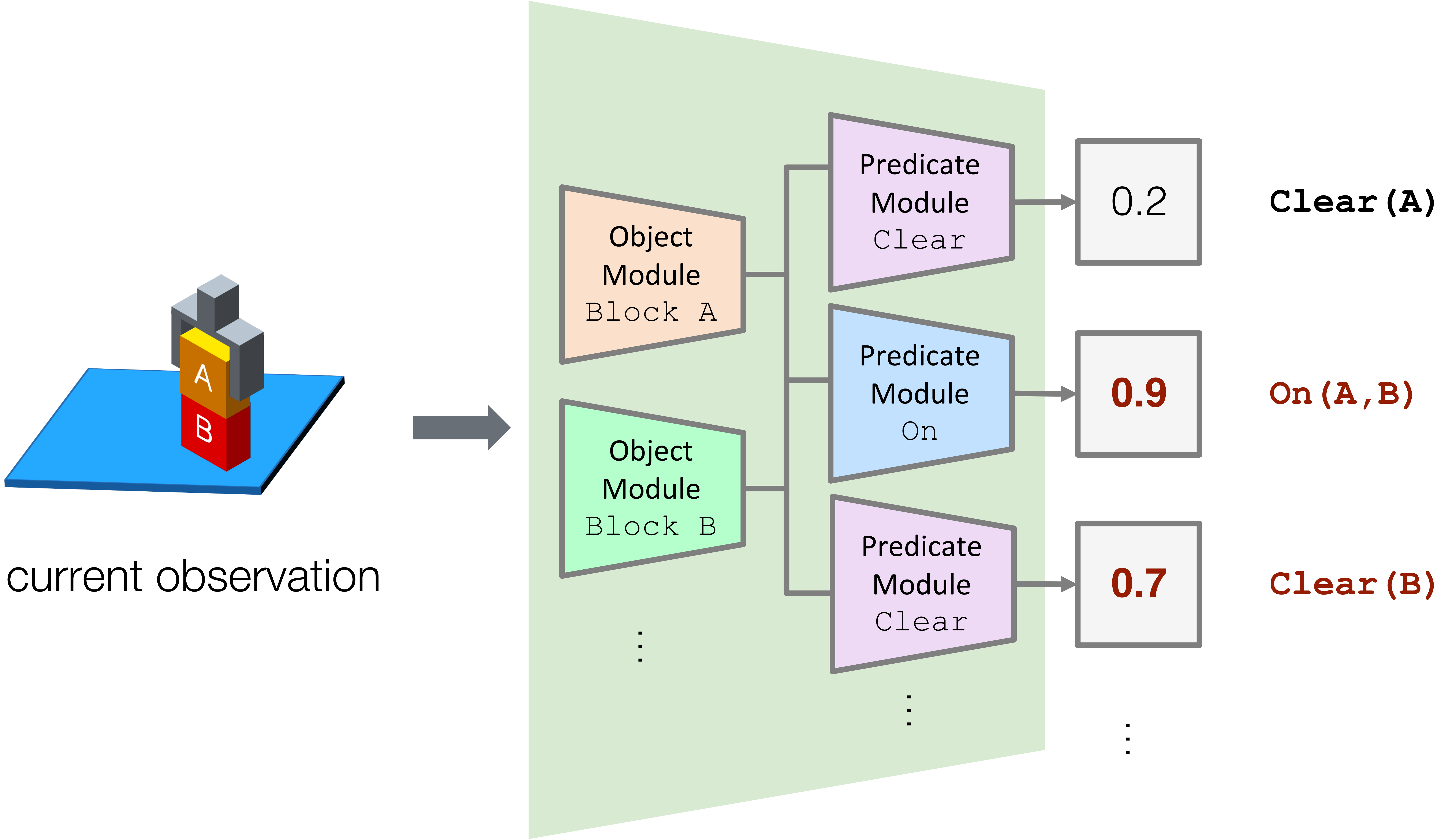
Huang et al. “Continuous Relaxation of Symbolic Planner”, IROS 2019

Planning-based One-Shot Visual Imitation: **Continuous Planner**



Huang et al. “Continuous Relaxation of Symbolic Planner”, IROS 2019

Planning-based One-Shot Visual Imitation: **Continuous Planner**



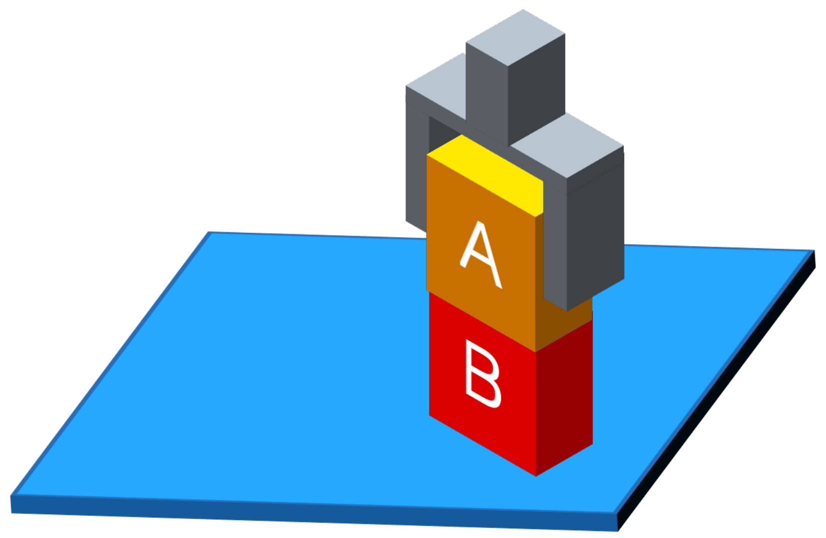
Predicate predictions are continuous arising from perceptual uncertainty.

Inconsistency might occur when we discretize the predictions.

Inconsistency leads to automatic planning failure (no way to reach the goal state).

Huang et al. "Continuous Relaxation of Symbolic Planner", IROS 2019

Planning-based One-Shot Visual Imitation: **Continuous Planner**



current observation

Symbol Grounding Networks

0.2	`Clear(A)`
0.9	`On(A,B)`
0.7	`Clear(B)`

```
(:Init
 (On A B)
 (Clear B))
```

discrete symbols

Concept definition or learning

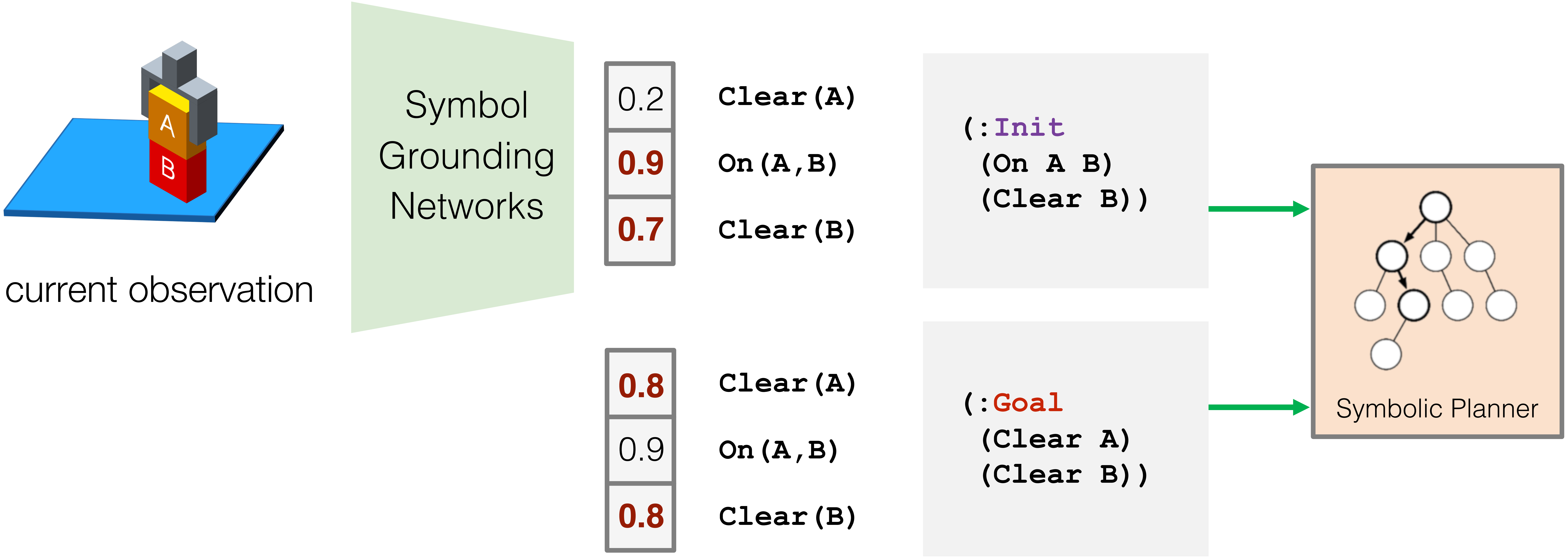
+ Markov Logic Networks

```
(clear (?block)
 :percepts  ((block ?block))
 :negatives ((on ?other?block)))
```

MAP inference

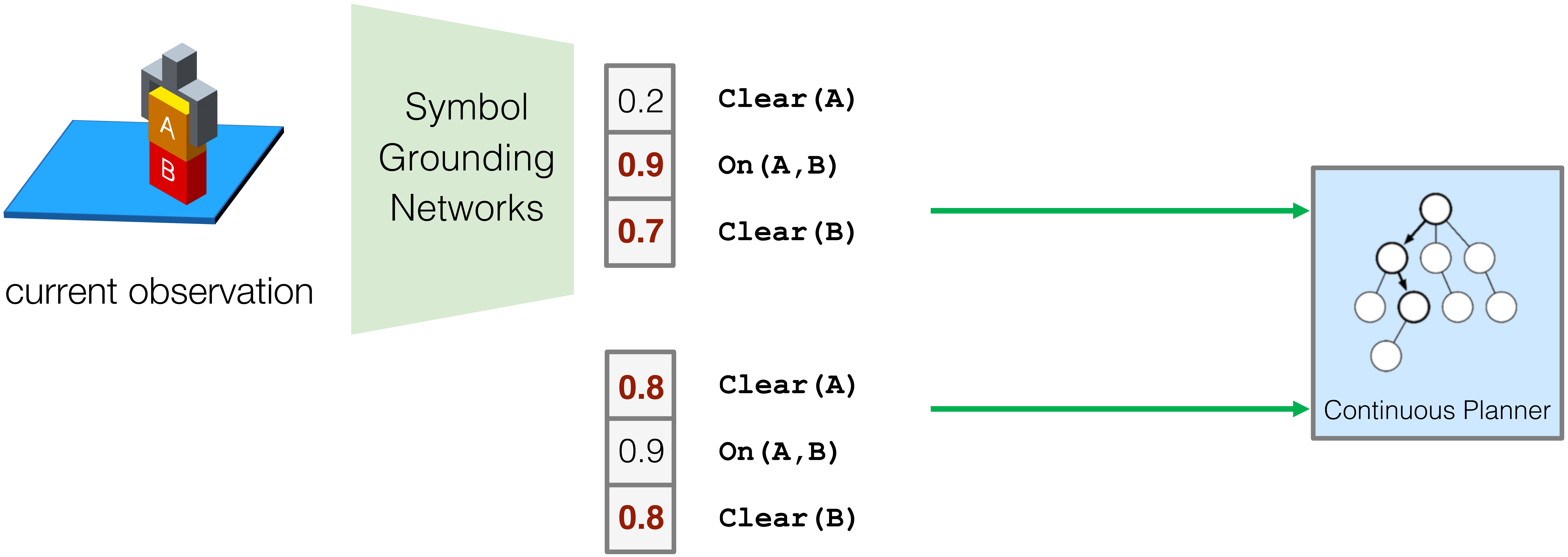
Huang et al. “Continuous Relaxation of Symbolic Planner”, IROS 2019

Planning-based One-Shot Visual Imitation: **Continuous Planner**



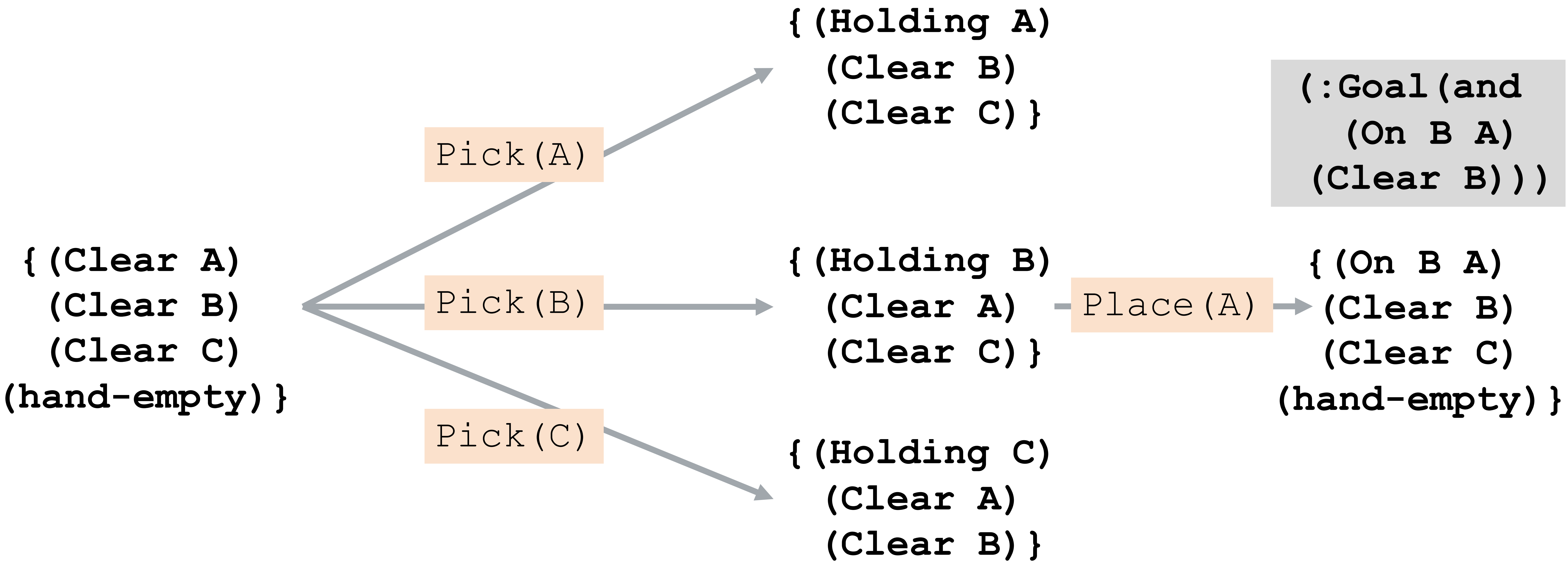
Huang et al. “Continuous Relaxation of Symbolic Planner”, IROS 2019

Planning-based One-Shot Visual Imitation: **Continuous Planner**



Huang et al. “Continuous Relaxation of Symbolic Planner”, IROS 2019

Planning-based One-Shot Visual Imitation: **Continuous Planner**



1. state representation

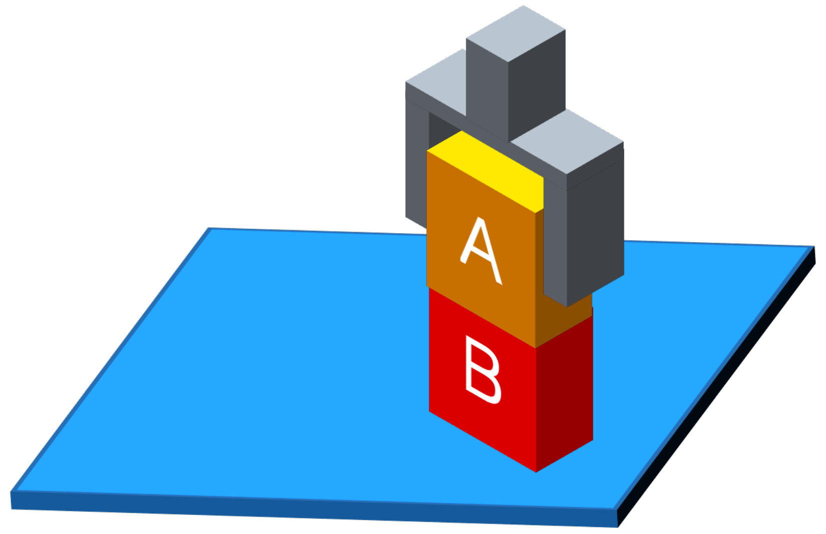
2. applicable actions

3. action application

4. goals satisfaction

Huang et al. "Continuous Relaxation of Symbolic Planner", IROS 2019

Planning-based One-Shot Visual Imitation: **Continuous Planner**



current observation

Symbol Grounding Networks

0.7	`Clear(A)`
0.9	`On(A,B)`
0.2	`Clear(B)`

`((On A B) (Clear A))` $0.7 \times 0.9 \times (1 - 0.2) = 0.7 \times 0.9 \times 0.2$

`((On A B) (Clear A) (Clear B))` $0.7 \times 0.9 \times 0.2$

1. state representation

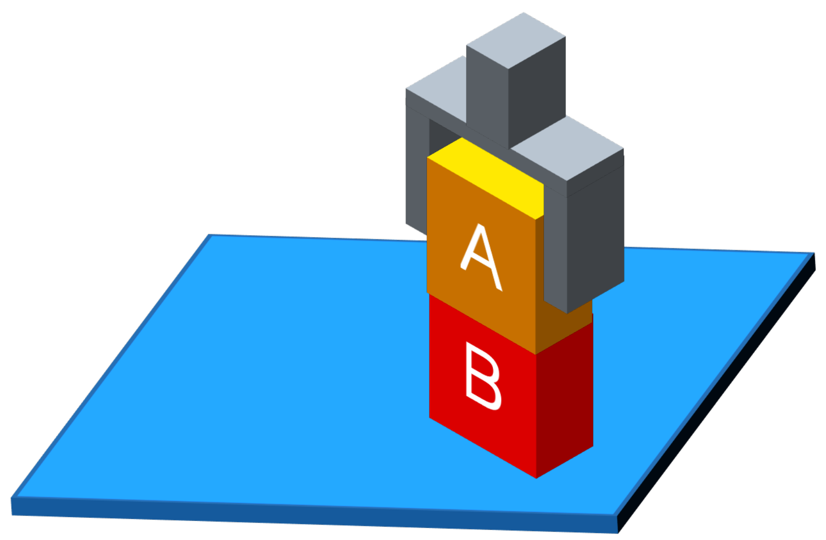
2. applicable actions

3. action application

4. goals satisfaction

Huang et al. “Continuous Relaxation of Symbolic Planner”, IROS 2019

Planning-based One-Shot Visual Imitation: **Continuous Planner**



current observation

```
Stack(B,A)

(:Precondition
 (and
  (Clear B)
  (Holding A)
   ...))
```

0.1	**Clear(A)**
0.9	**Holding(A)**
0.6	**Clear(B)**

$0.6 \times 0.9 \times \ \ldots$

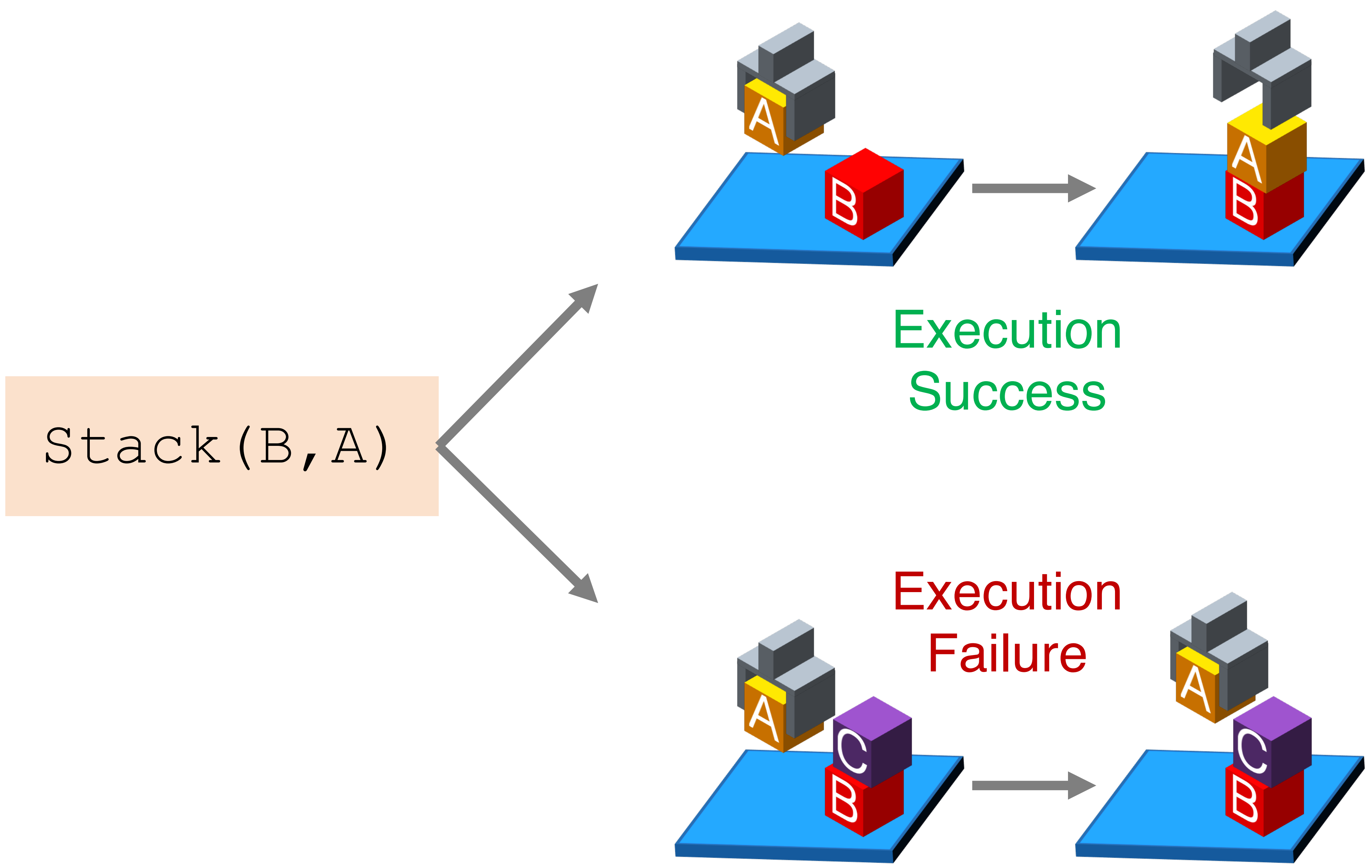
1. state representation	2. applicable actions	3. action application	4. goals satisfaction

Huang et al. "Continuous Relaxation of Symbolic Planner", IROS 2019

Planning-based One-Shot Visual Imitation: **Continuous Planner**

0.1	`Clear(A)`
0.9	**`Holding(A)`**
0.6	**`Clear(B)`**

$0.6 \times 0.9 \times$ ...



1. state representation

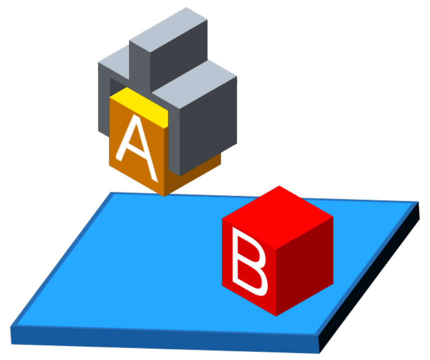
2. applicable actions

3. action application

4. goals satisfaction

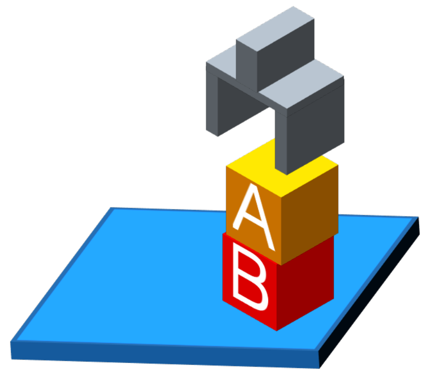
Huang et al. "Continuous Relaxation of Symbolic Planner", IROS 2019

Planning-based One-Shot Visual Imitation: **Continuous Planner**



current state distribution

0.8	**On(B,A)**
0.9	**Clear(B)**
0.2	**Clear(A)**



goal state distribution

0.9	**On(B,A)**
0.8	**Clear(B)**
0.5	**Clear(A)**

1. state representation

2. applicable actions

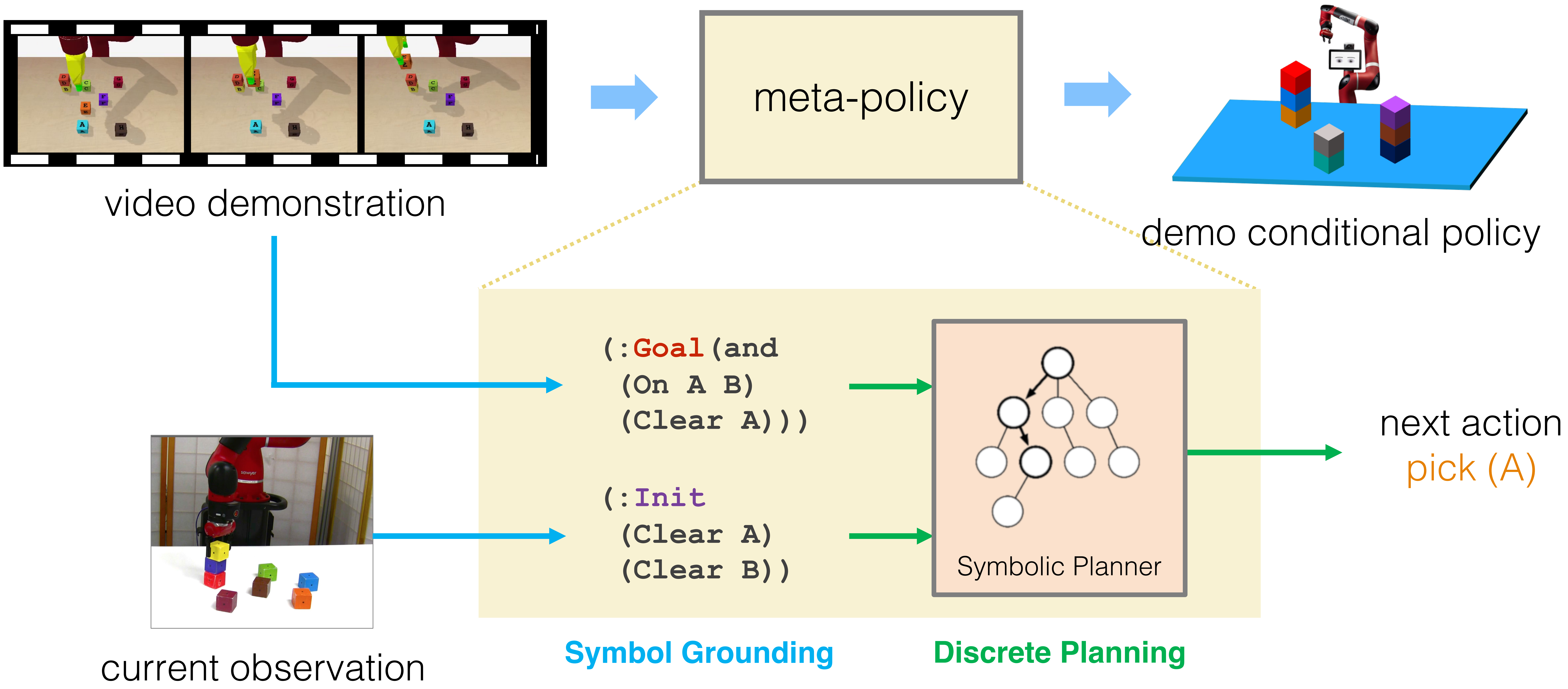
3. action application

4. goals satisfaction

Huang et al. "Continuous Relaxation of Symbolic Planner", IROS 2019

Planning-based One-Shot Visual Imitation: **Continuous Planner**

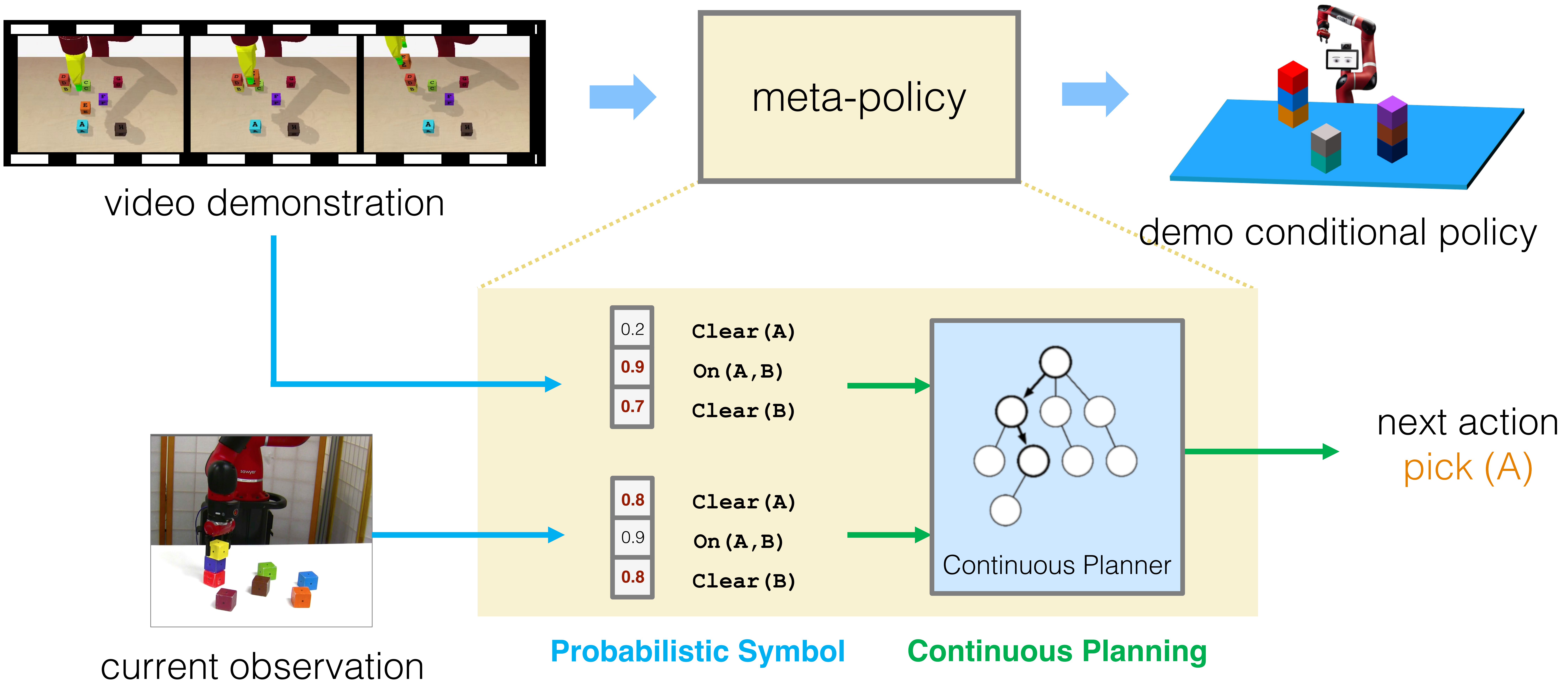
Key idea: Continuous Relaxation of Discrete Symbolic Reasoning



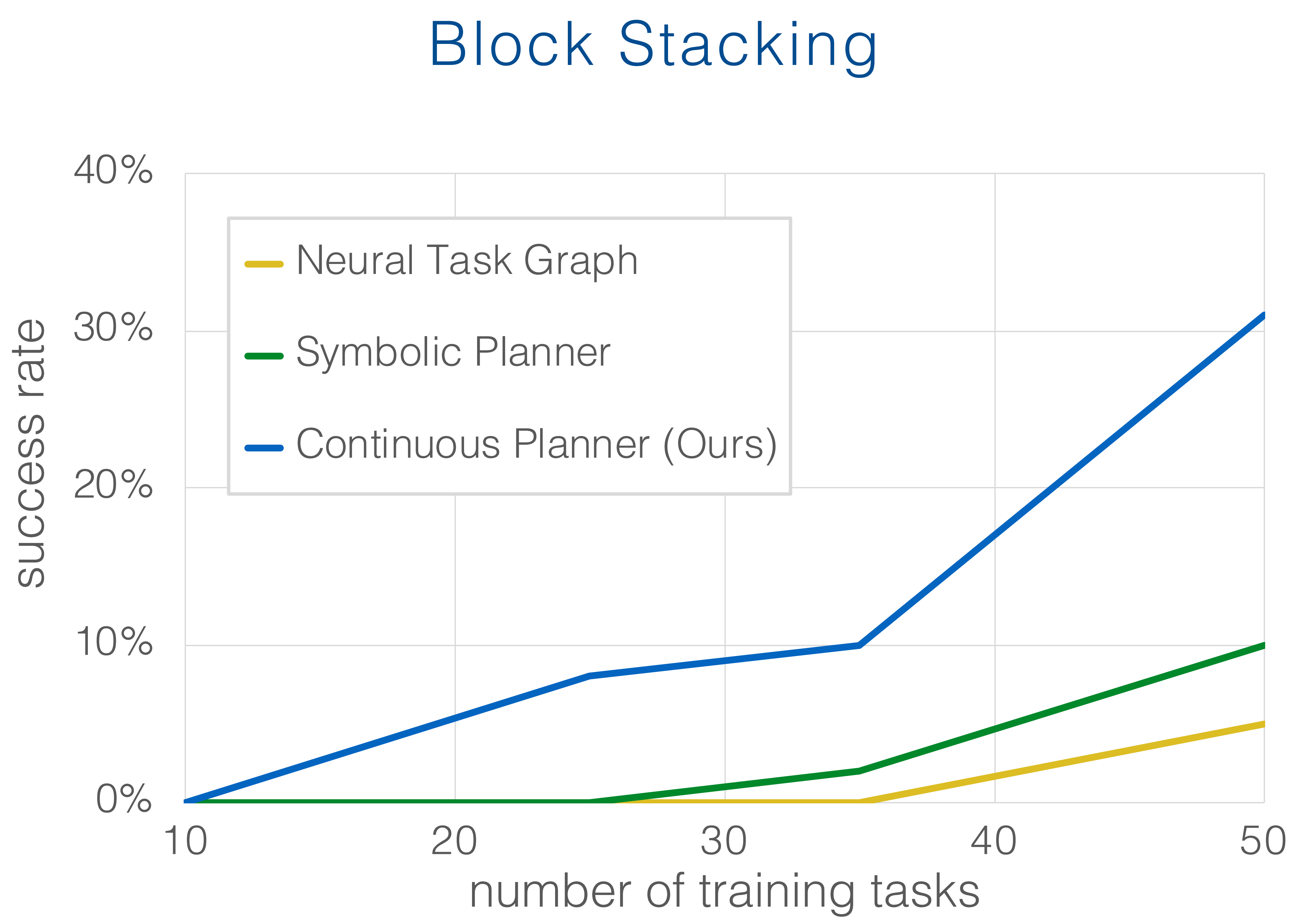
Huang et al. “Continuous Relaxation of Symbolic Planner”, IROS 2019

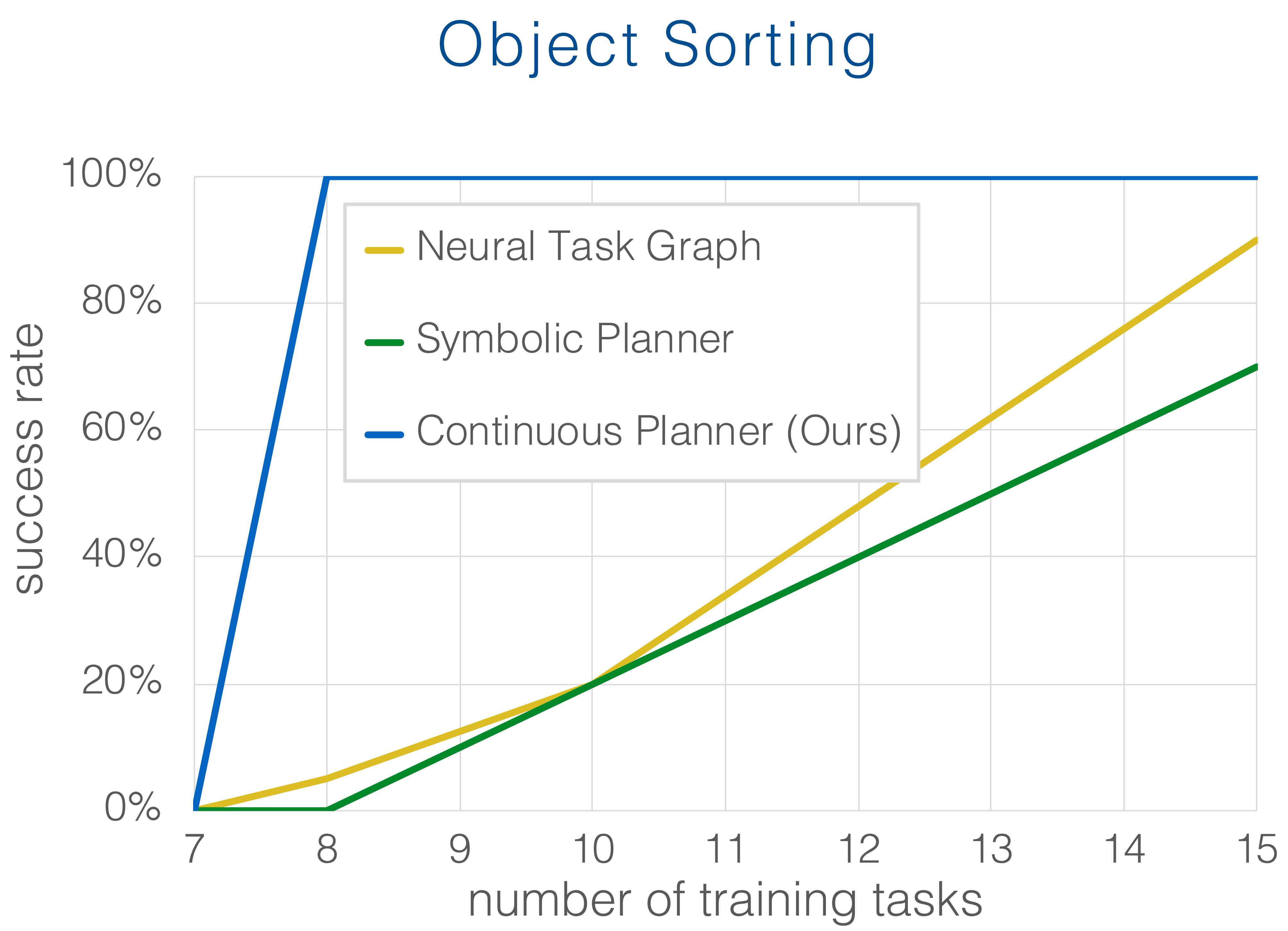
Planning-based One-Shot Visual Imitation: **Continuous Planner**

Key idea: Continuous Relaxation of Discrete Symbolic Reasoning



Planning-based One-Shot Visual Imitation: **Continuous Planner**



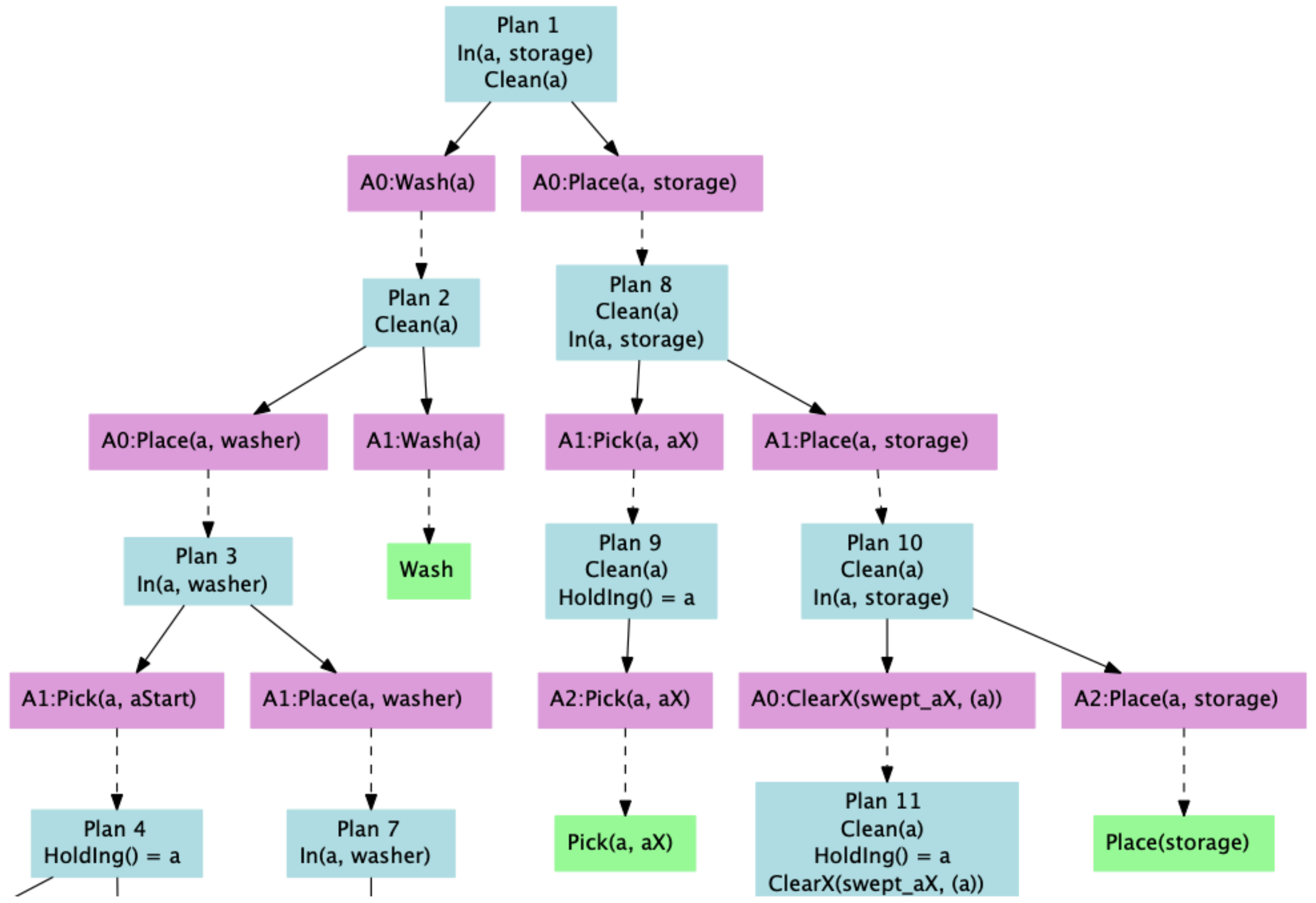


Stronger generalization and alternative execution orders from limited training demonstrations

Huang et al. "Continuous Relaxation of Symbolic Planner", IROS 2019

Visual Imitation Learning

integrating **deep learning** with **symbolic reasoning**

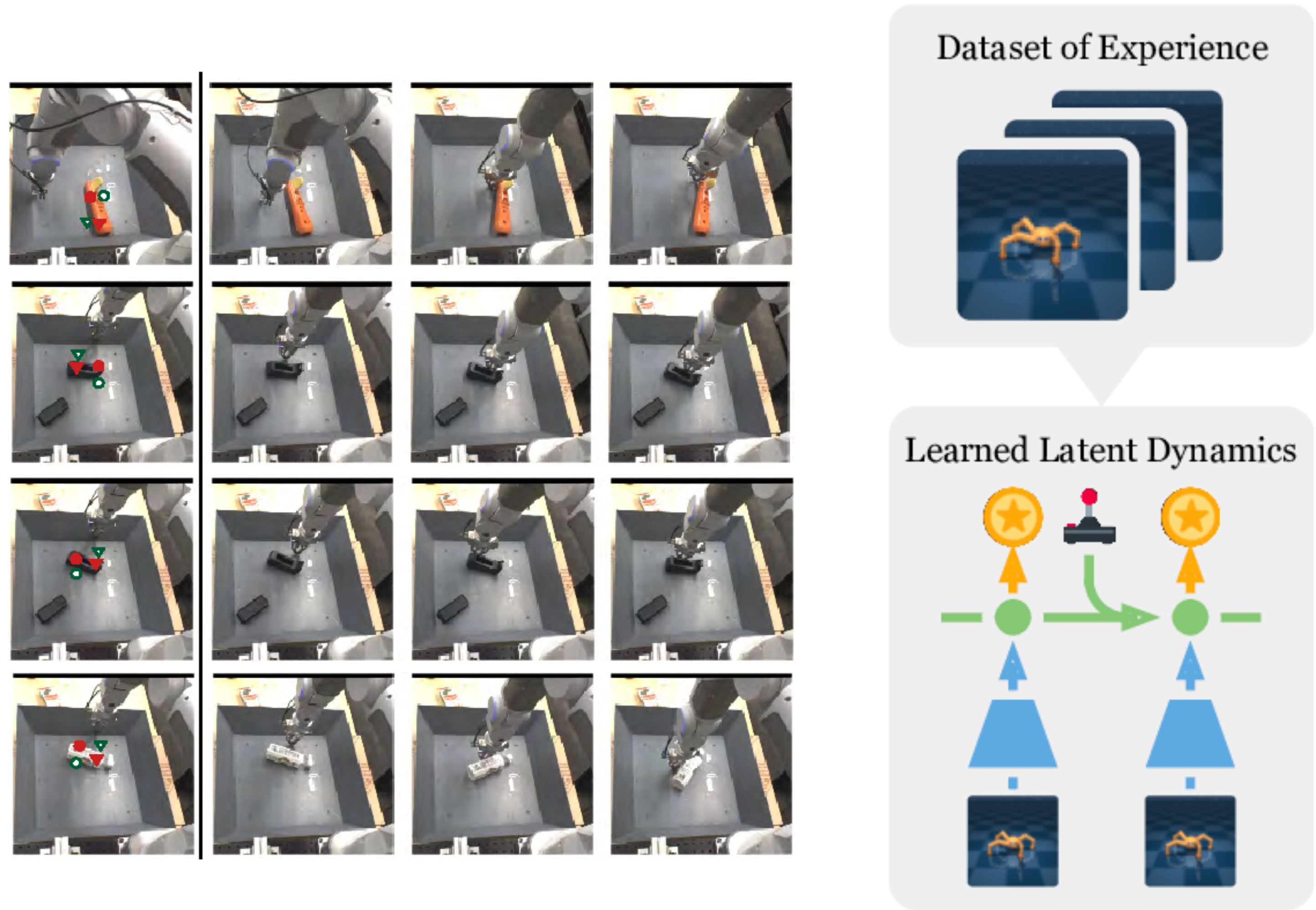


[Waldinger 1975; Korf 1987; Kaelbling ICRA'11]

classical symbolic planning

human-interpretable and long-horizon

symbols and planning domain required



[Finn et al. ICRA'17; Oh et al. NIPS'15; Hafner et al. ICLR'20]

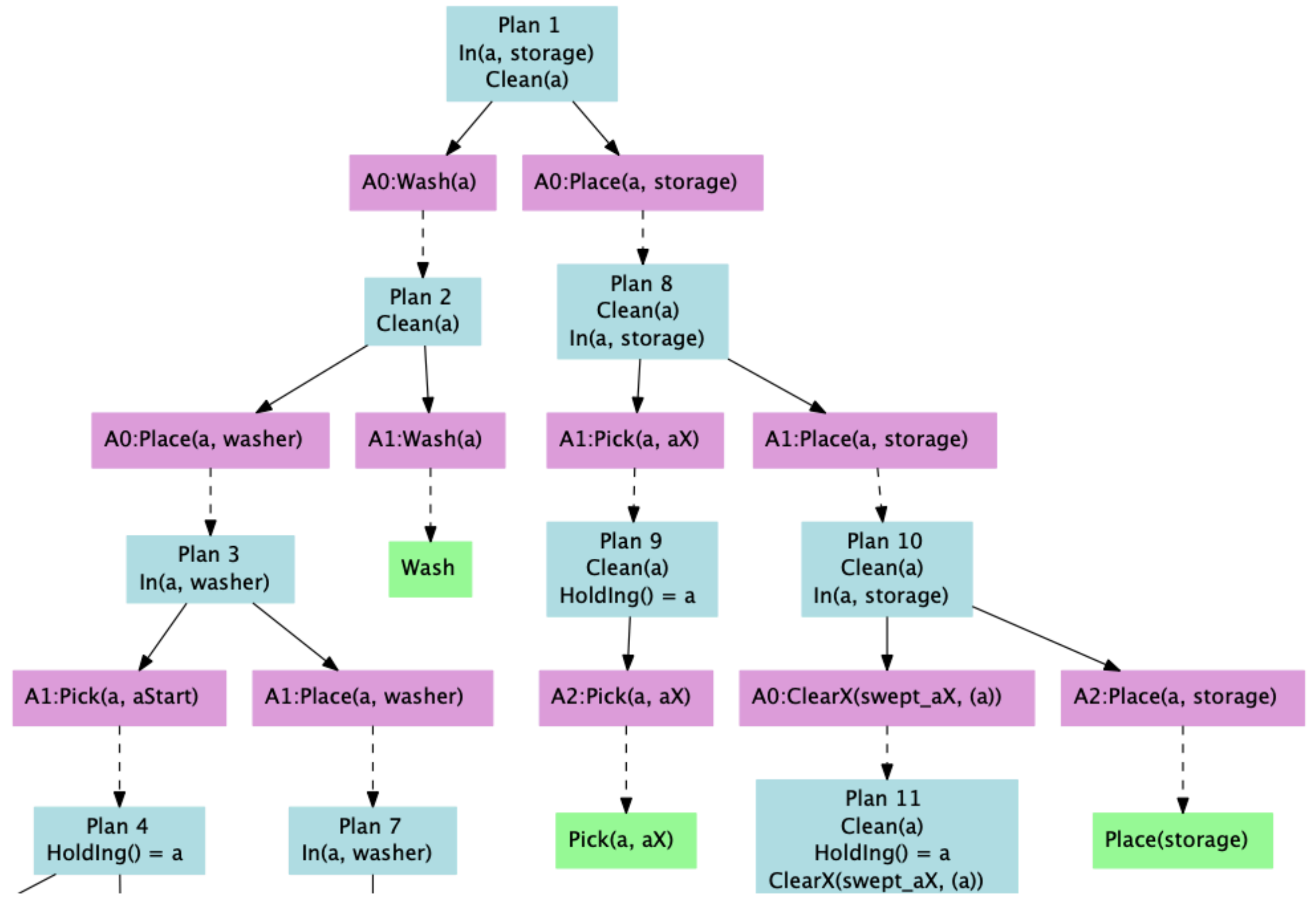
plan from observations

grounded on raw sensory data

myopic sampling, short-horizon tasks

Visual Imitation Learning

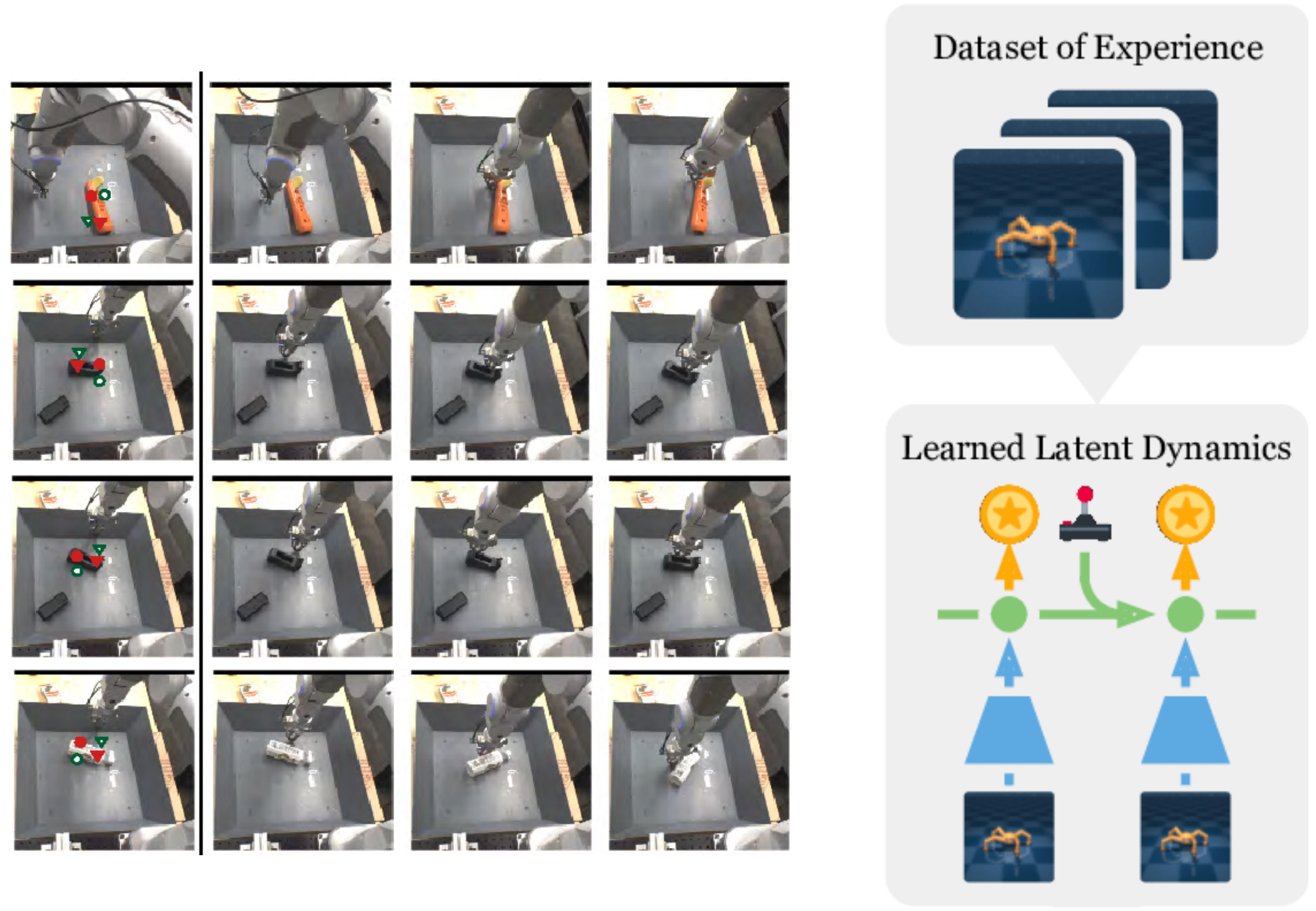
integrating **deep learning** with **symbolic reasoning**



[Waldinger 1975; Korf 1987; Kaelbling ICRA’11]

classical symbolic planning

+



[Finn et al. ICRA’17; Oh et al. NIPS'15; Hafner et al. ICLR’20]

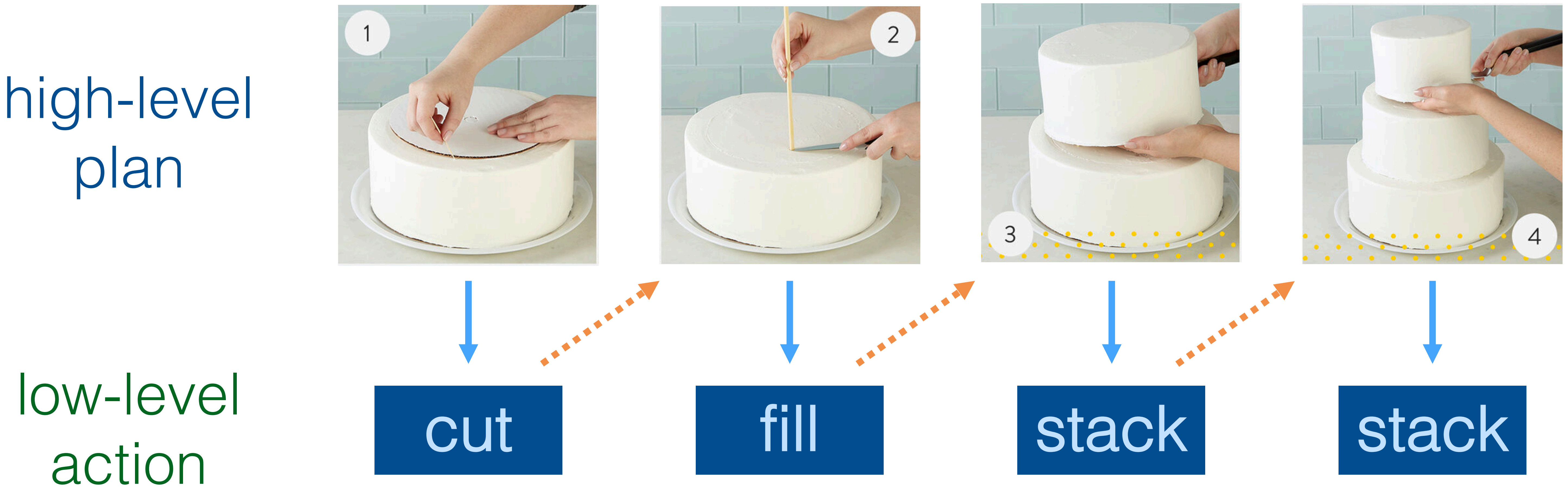
plan from observations

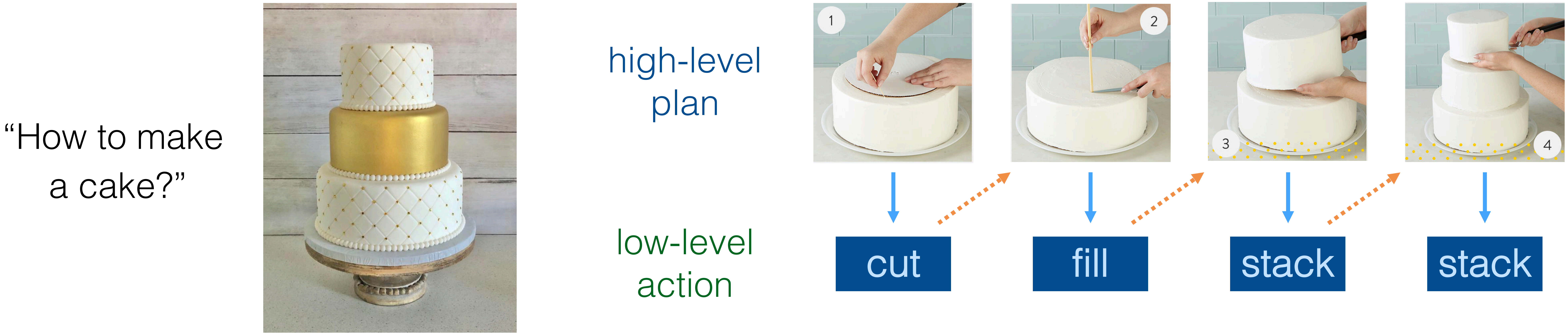
plan backward (regression planning) in a symbolic space conditioning on the visual observation

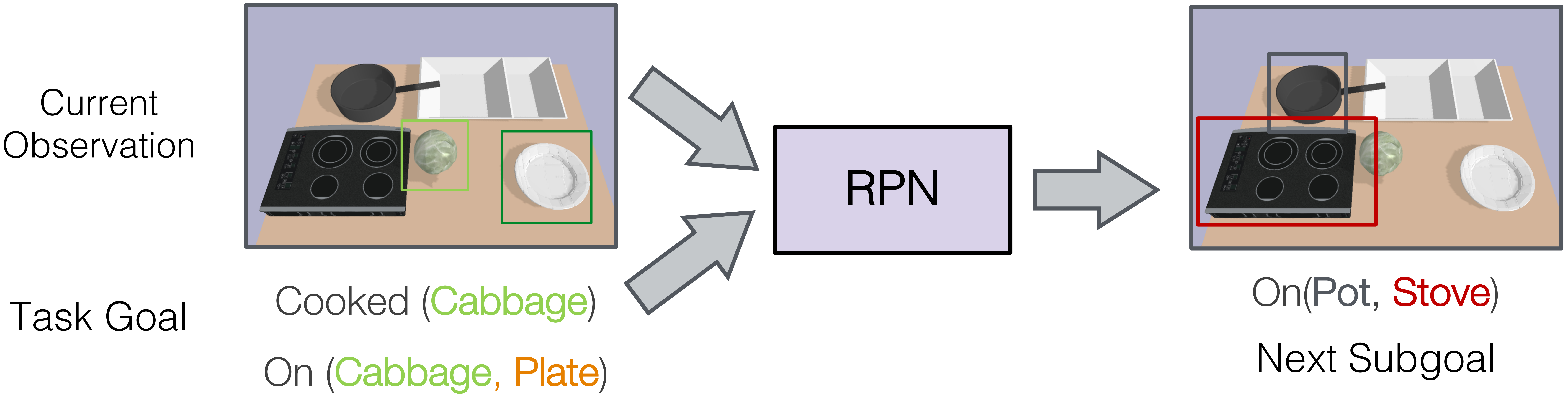
No planning domain required

Strong generalization to long-horizon tasks

“How to make a cake?”

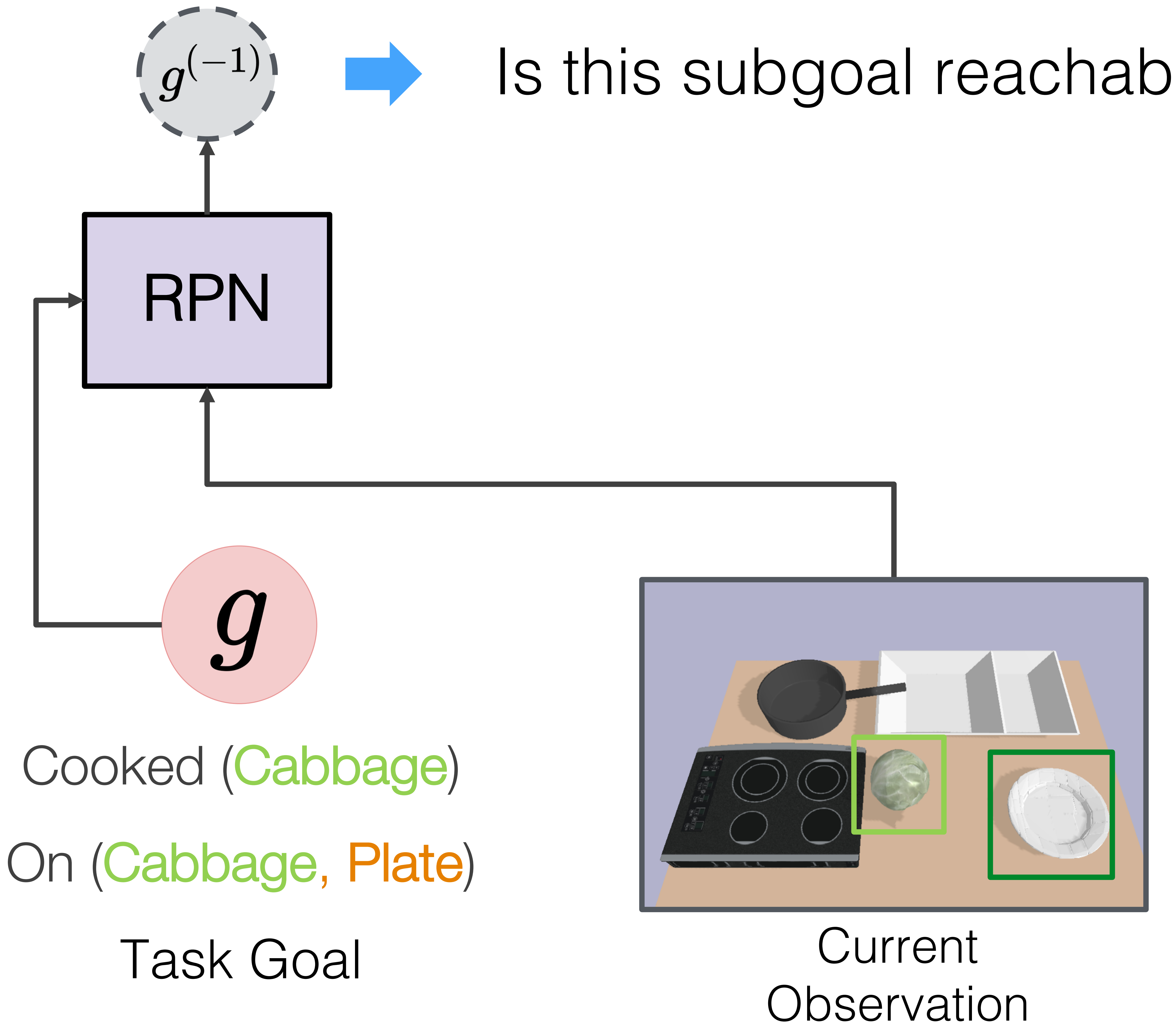



"How to make a cake?"
high-level plan
low-level action
1
2
3
4
cut
fill
stack
stack

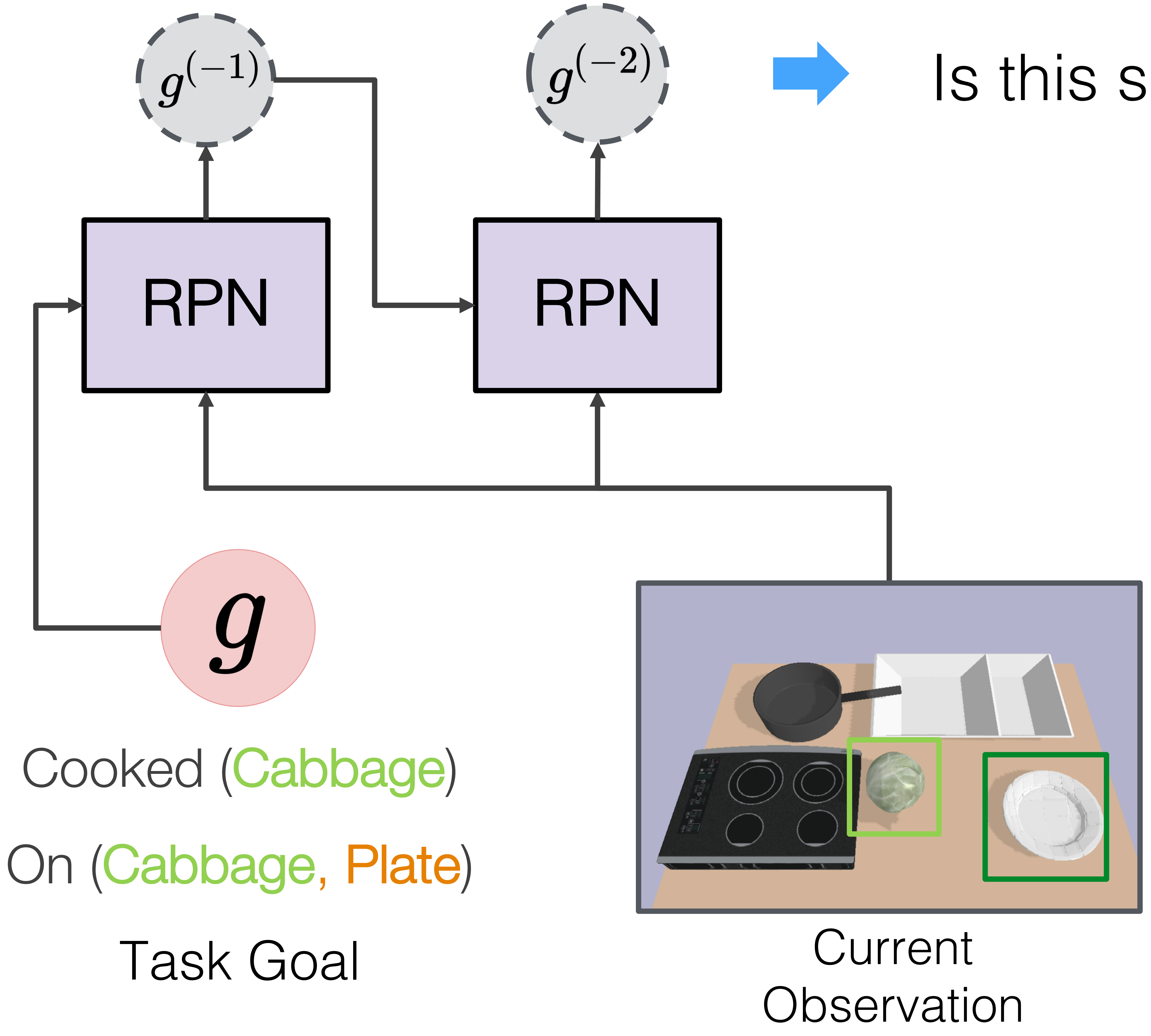

Current Observation
Task Goal
Cooked (Cabbage)
On (Cabbage, Plate)
RPN
On(Pot, Stove)
Next Subgoal

Learning to Plan from Observations: **Regression Planning Network**

Is this subgoal reachable? **NO!**

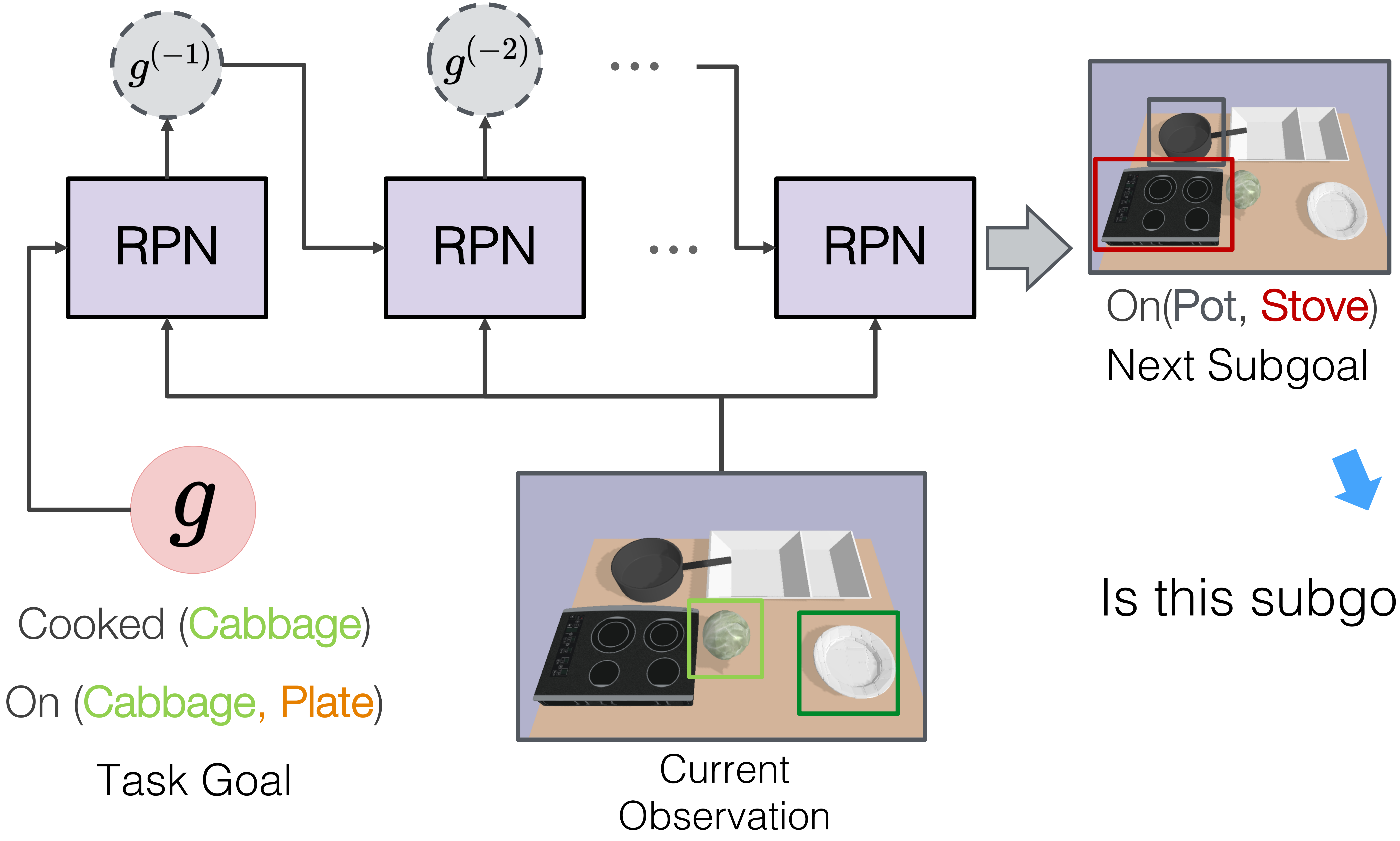


Learning to Plan from Observations: **Regression Planning Network**



Is this subgoal reachable? **NO!**

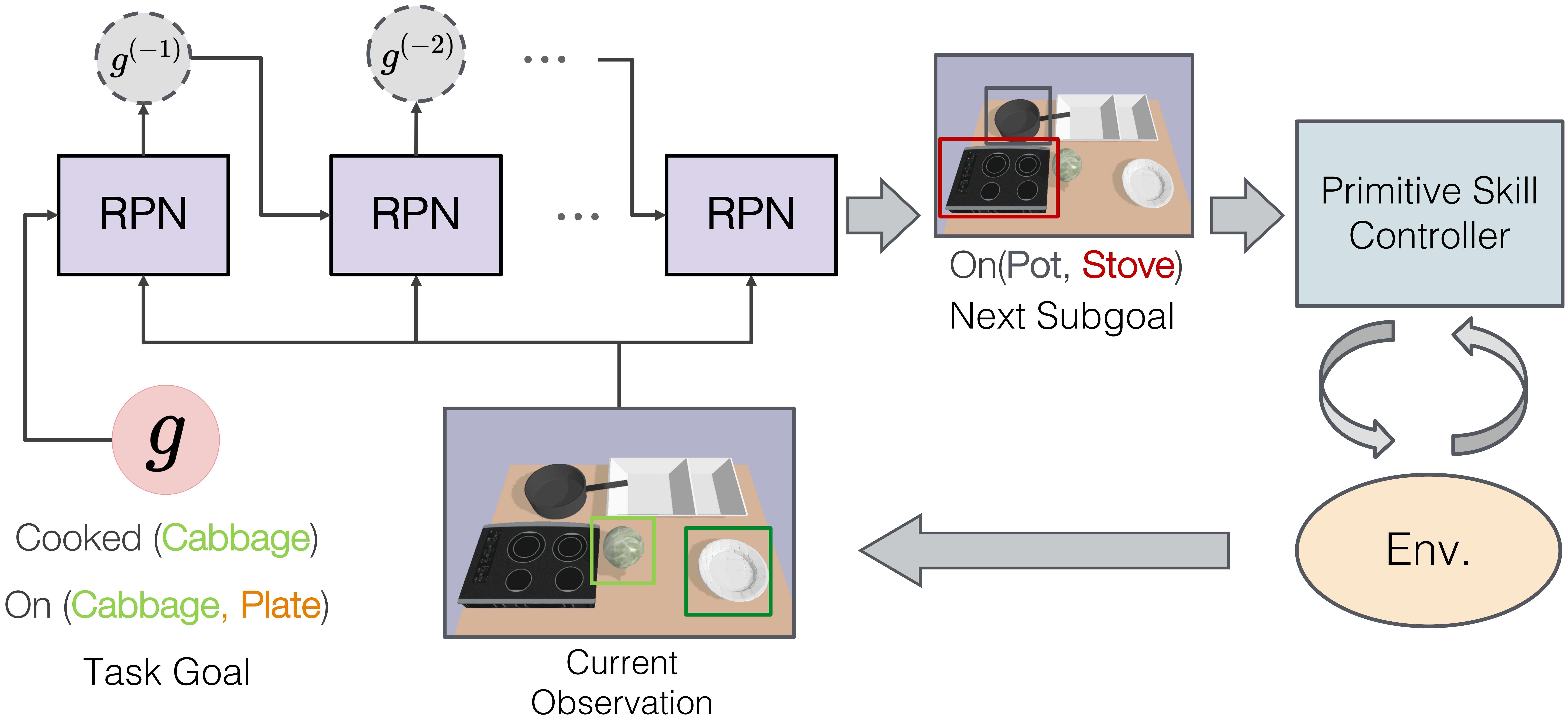
Learning to Plan from Observations: **Regression Planning Network**



Is this subgoal reachable? **YES!**

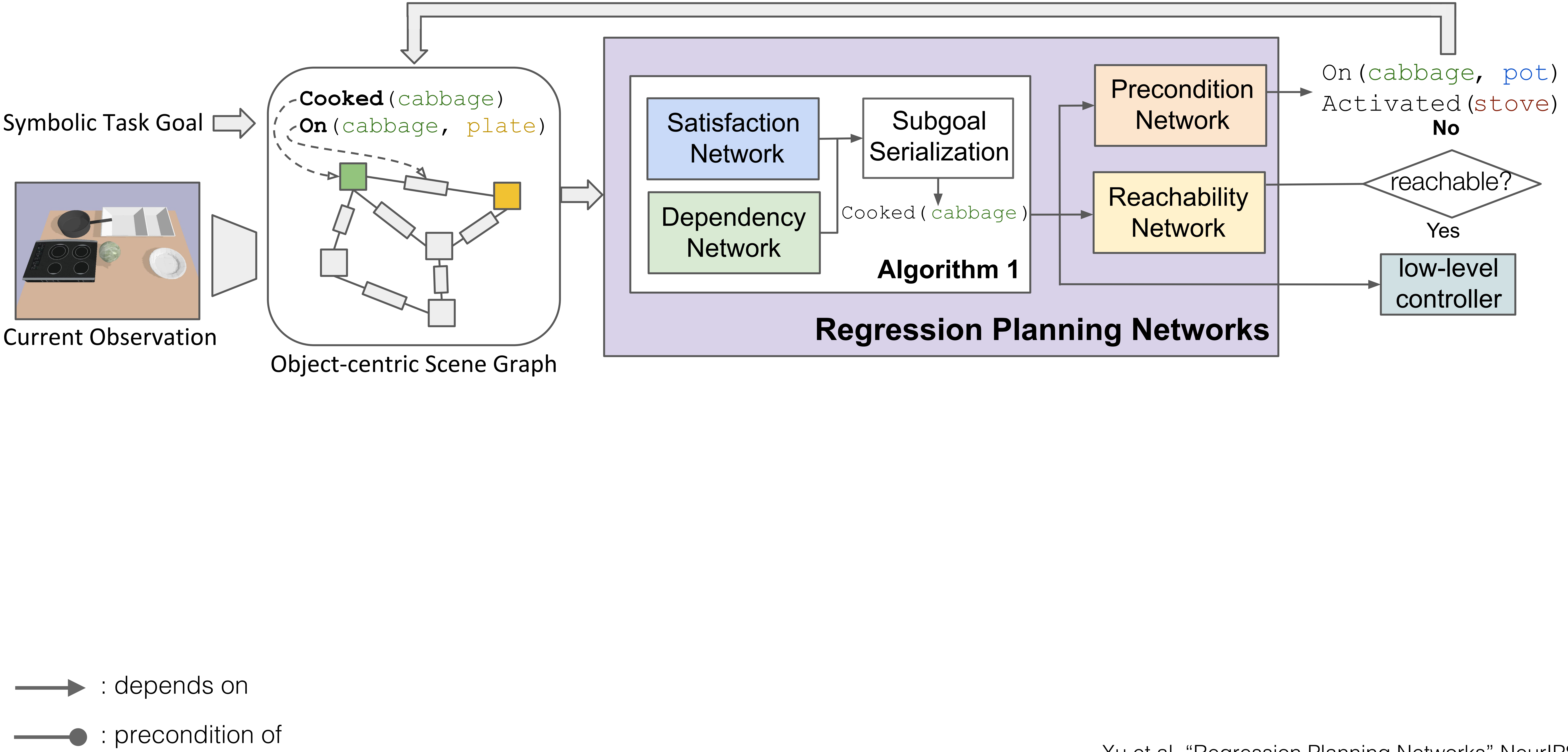
Xu et al. "Regression Planning Networks" NeurIPS 2019

Learning to Plan from Observations: **Regression Planning Network**



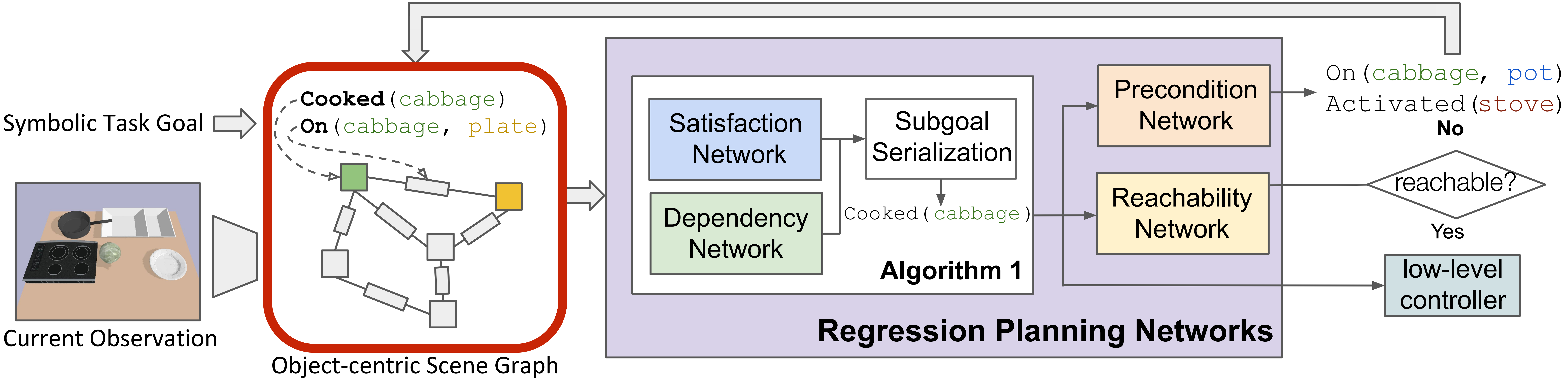
Xu et al. "Regression Planning Networks" NeurIPS 2019

Learning to Plan from Observations: Regression Planning Network



Xu et al. "Regression Planning Networks" NeurIPS 2019

Learning to Plan from Observations: **Regression Planning Network**



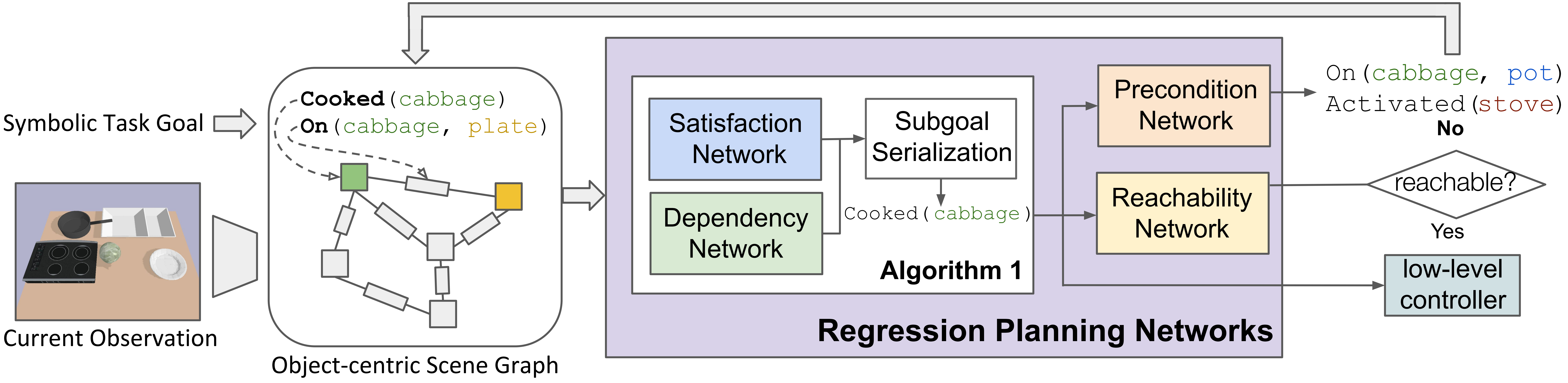
Scene graph as **object-centric representations** for entities and relationships

: depends on

: precondition of

Xu et al. “Regression Planning Networks” NeurIPS 2019

Learning to Plan from Observations: Regression Planning Network



Cooked(cabbage)	
Sat: **False**	Rec: **False**

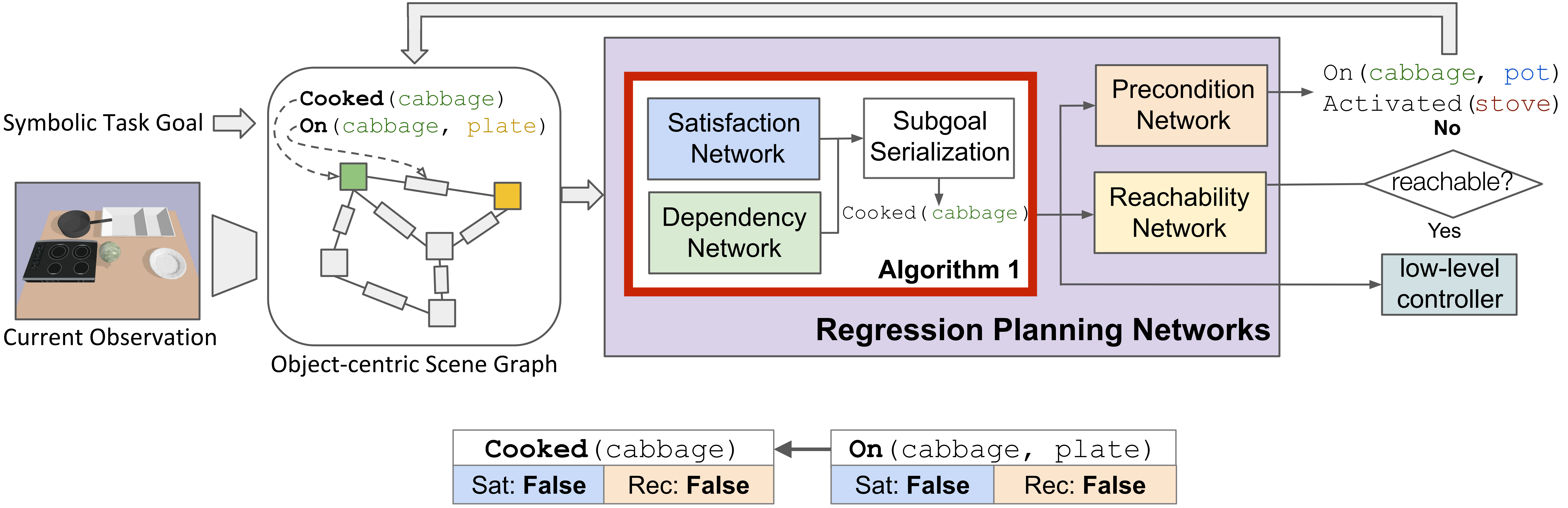
On(cabbage, plate)	
Sat: **False**	Rec: **False**

: depends on

: precondition of

Xu et al. "Regression Planning Networks" NeurIPS 2019

Learning to Plan from Observations: Regression Planning Network



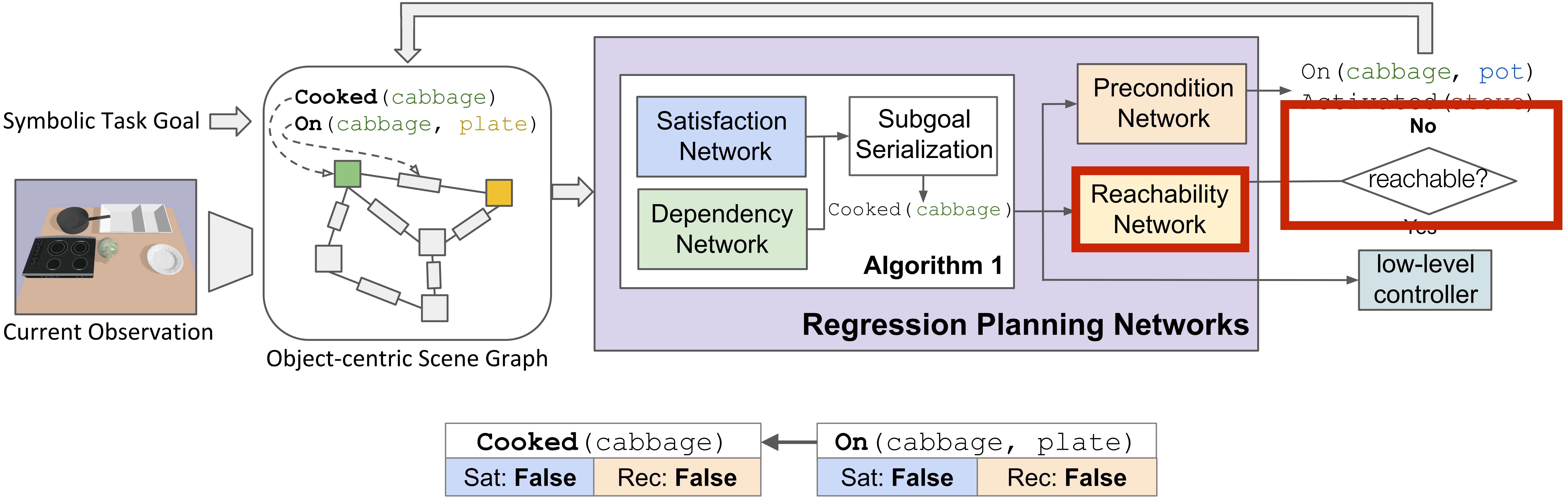
Subgoal serialization: determine the execution order of subgoals by predicting their dependencies

: depends on

: precondition of

Xu et al. "Regression Planning Networks" NeurIPS 2019

Learning to Plan from Observations: **Regression Planning Network**



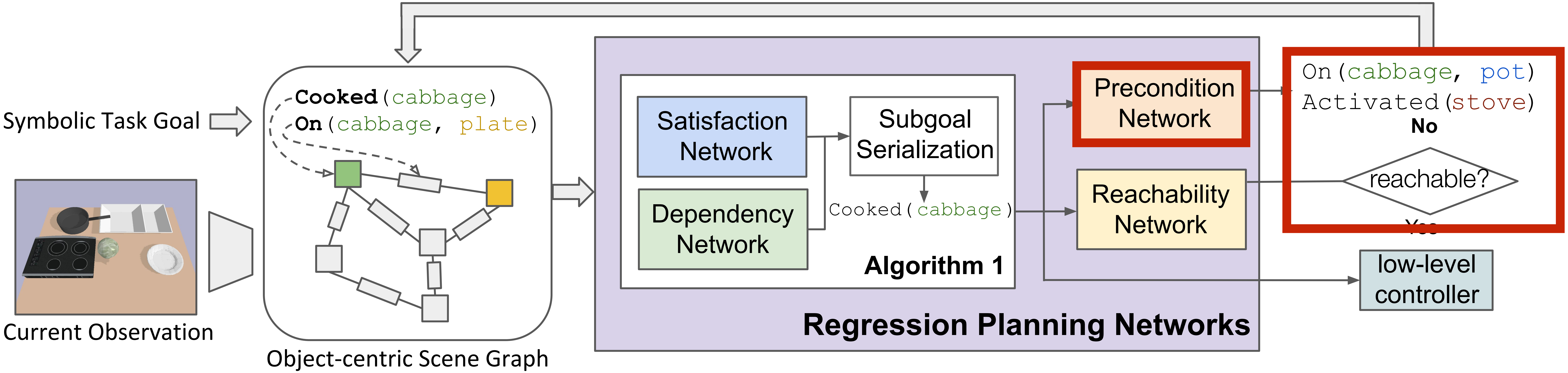
Reachability: determines if the subgoal can be achieved by a low-level primitive skill

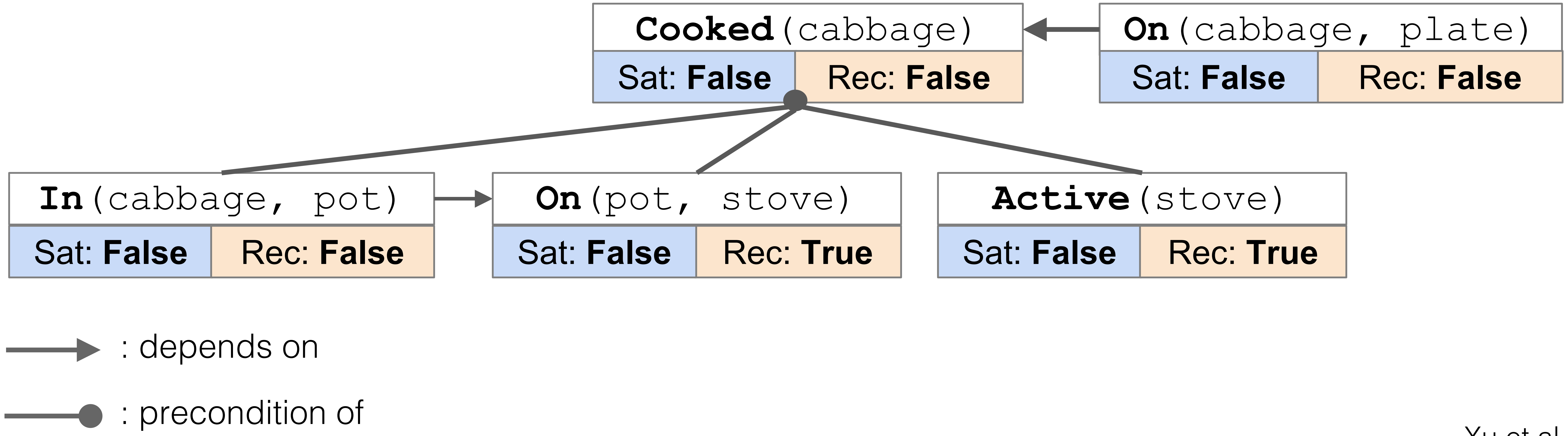
: depends on

: precondition of

Xu et al. “Regression Planning Networks” NeurIPS 2019

Learning to Plan from Observations: Regression Planning Network

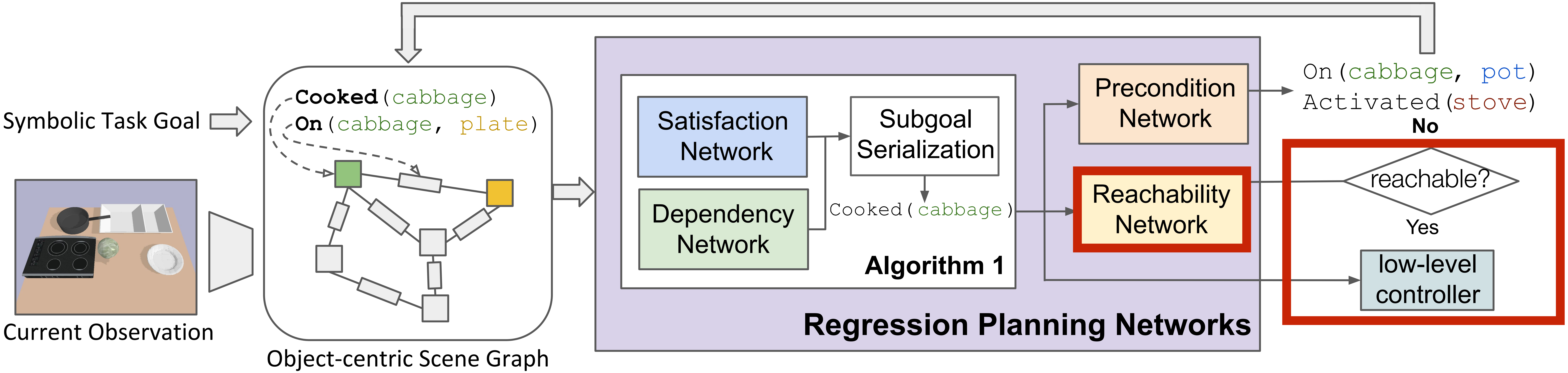


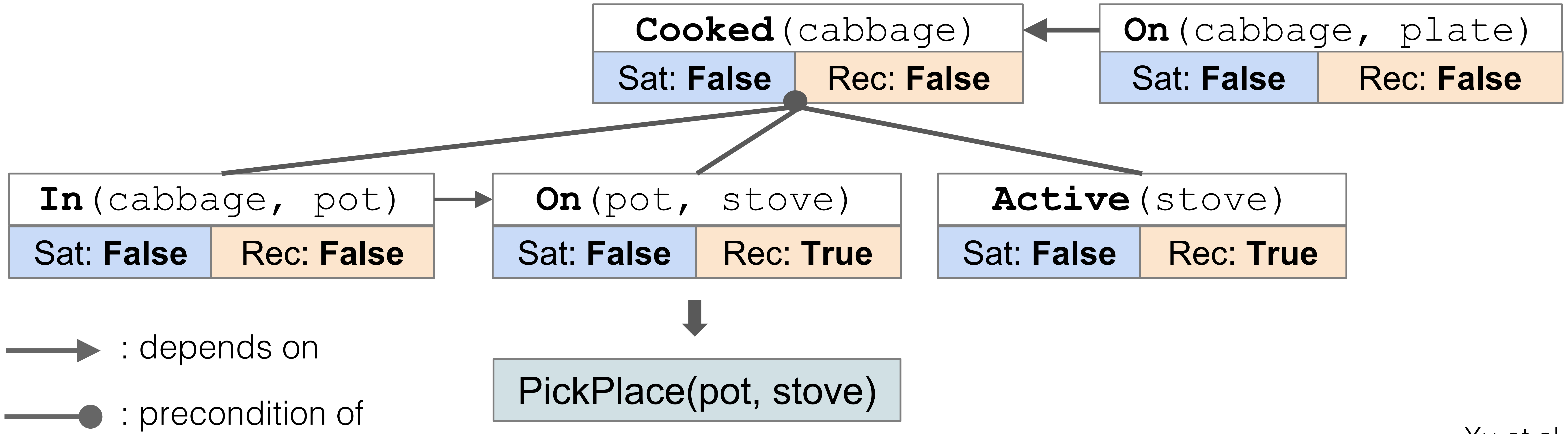


Precondition network: Predict the predecessor subgoals that need to be satisfied

Xu et al. “Regression Planning Networks” NeurIPS 2019

Learning to Plan from Observations: **Regression Planning Network**



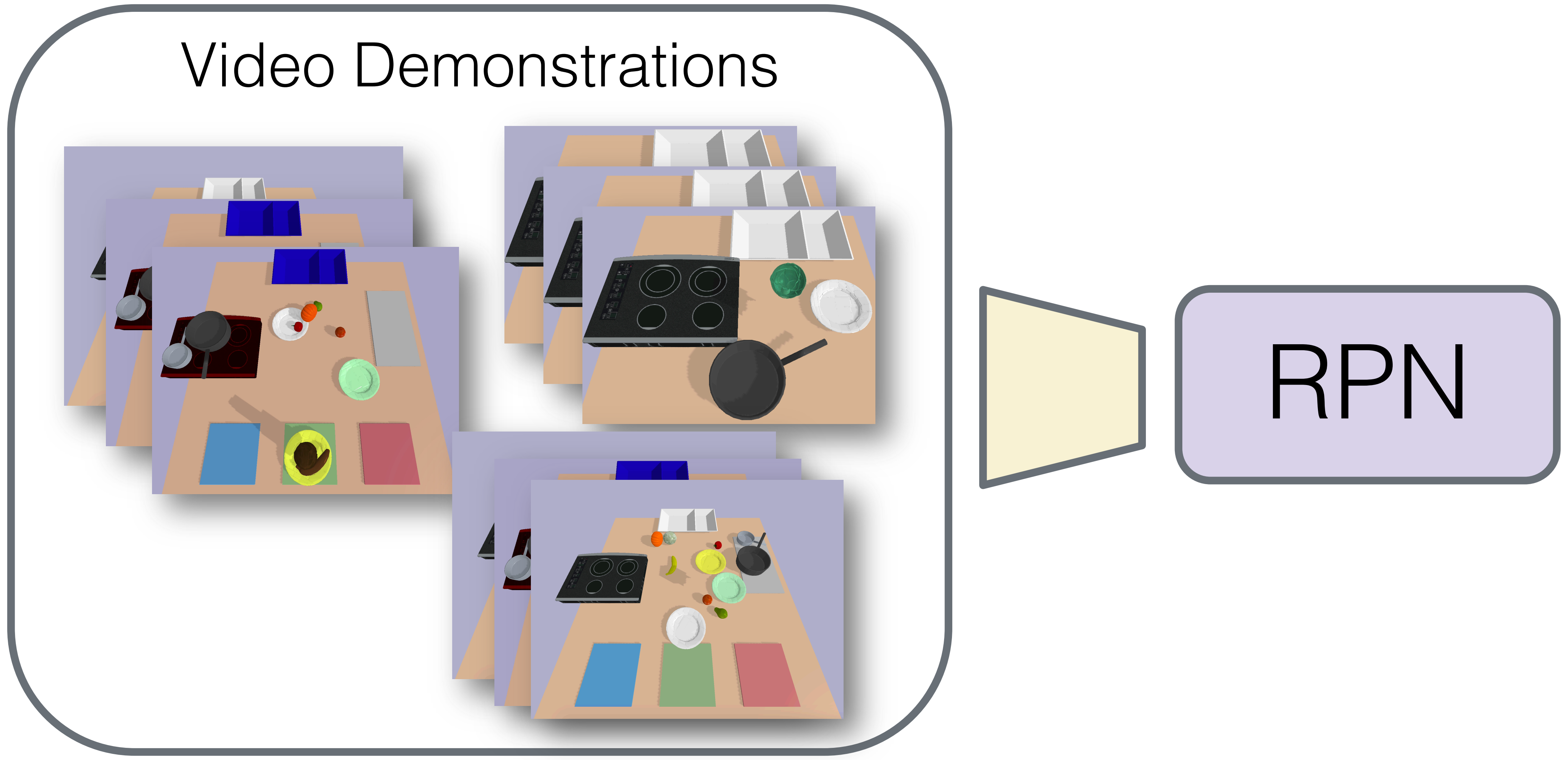


Low-level controller: Execute skills to achieve the subgoal

Xu et al. "Regression Planning Networks" NeurIPS 2019

Learning to Plan from Observations: **Regression Planning Network**

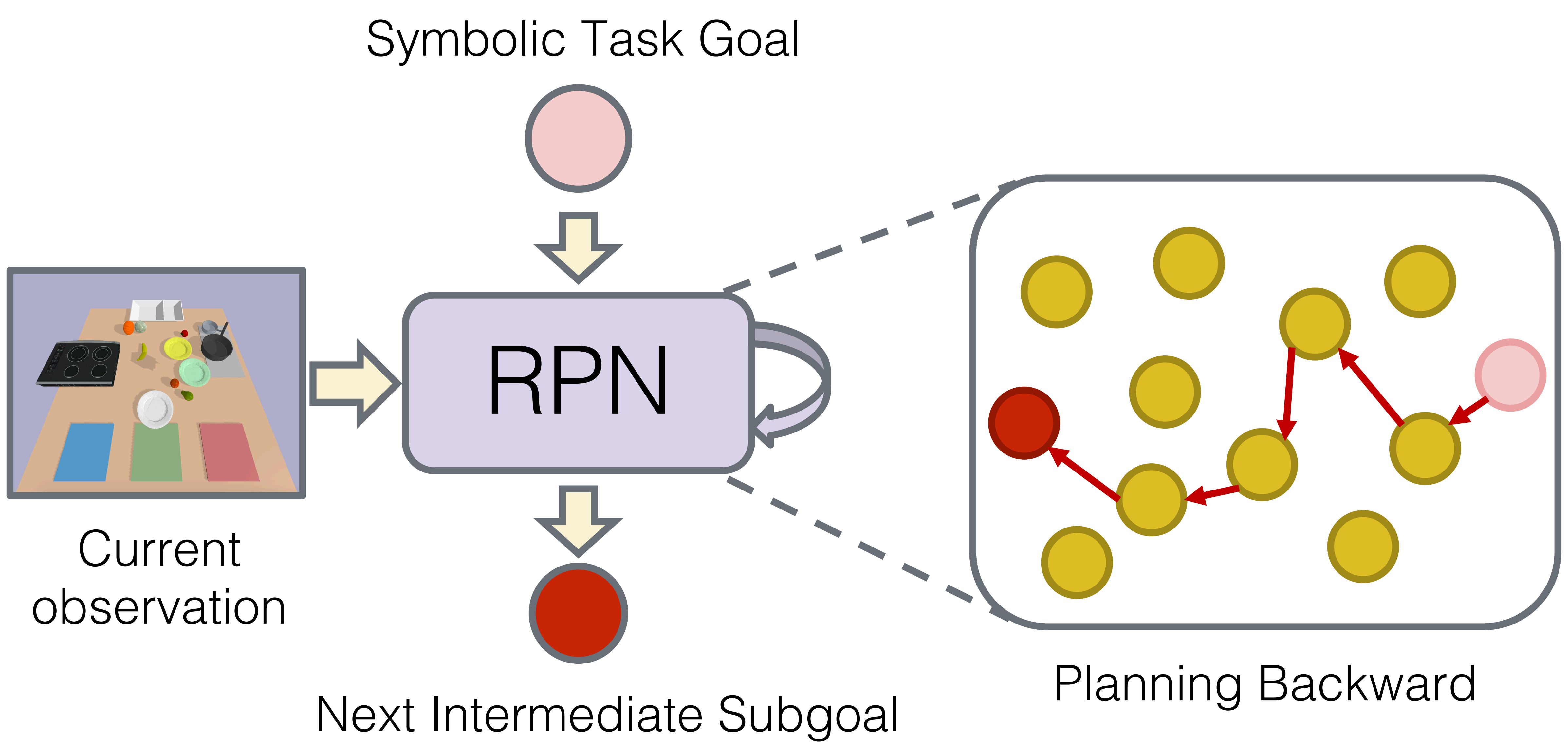
Learning



Trained on expert demonstrations

{(observation o_t, action a_t, subgoal g_t)}

Planning



Tested on unseen task goals

(does not require test-time demonstration)

Learning to Plan from Observations: Regression Planning Network

Goal Multiplicity (DoorKey)

Train: Open two doors

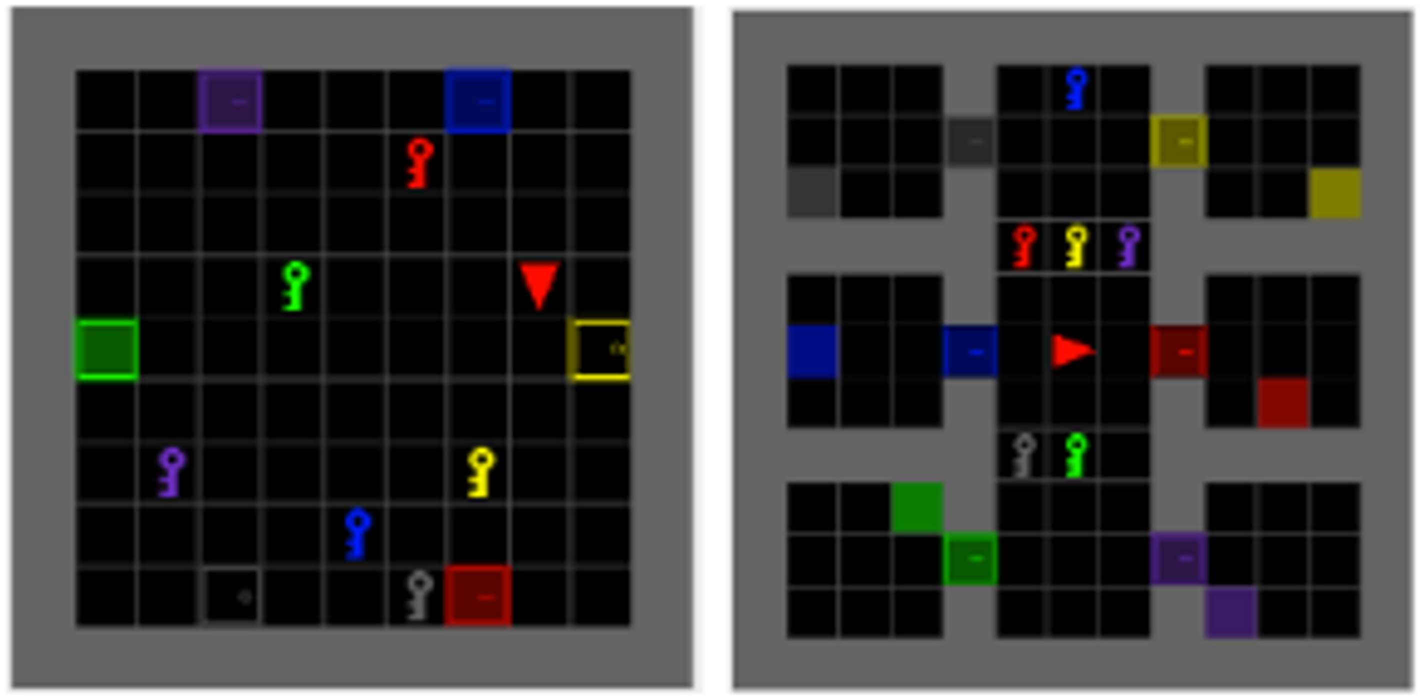
Test: Open $D > 2$ doors

Plan Composition (RoomGoal)

Train: 1. get key → open door

2. open door → reach goal

Test: get key → open door → reach goal



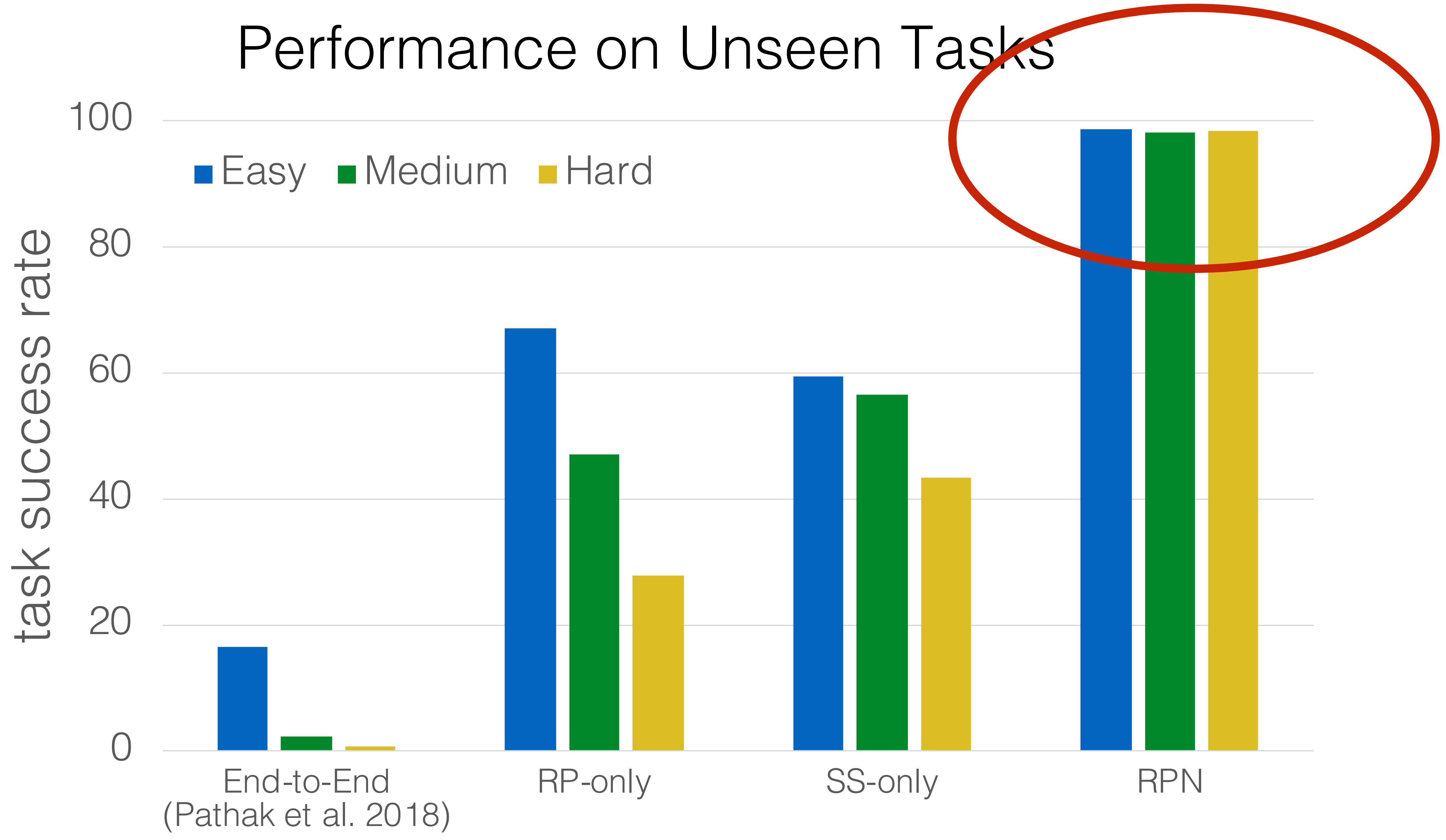
Domain	DoorKey			RoomGoal		
	Train	Eval		Train		Eval
Task	D=2	D=4	D=6	k-d	d-g	k-d-g
E2E [34]	81.2	1.2	0.0	100.0	100.0	3.2
RP-only	92.2	18.2	0.0	100.0	100.0	**100.0**
SS-only	99.7	46.0	21.1	99.9	100.0	7.8
RPN	99.1	**91.9**	**64.3**	98.7	99.9	**98.8**

Xu et al. "Regression Planning Networks" NeurIPS 2019

Learning to Plan from Observations: **Regression Planning Network**

near perfect generalization to longer tasks

Model trained on **2** dishes with **3** ingredients



Qualitative

(cook **3** dishes with **4** ingredients)

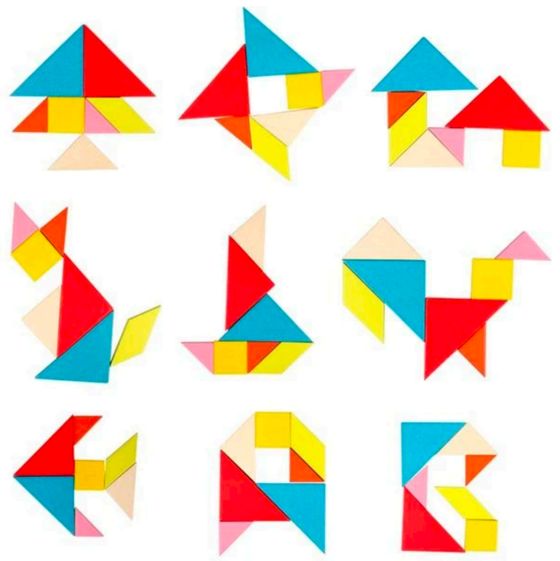
Quantitative

(the higher the better)

Xu et al. "Regression Planning Networks" NeurIPS 2019

Conclusions

Visual imitation Learning from video demonstrations



Compositional Generalization

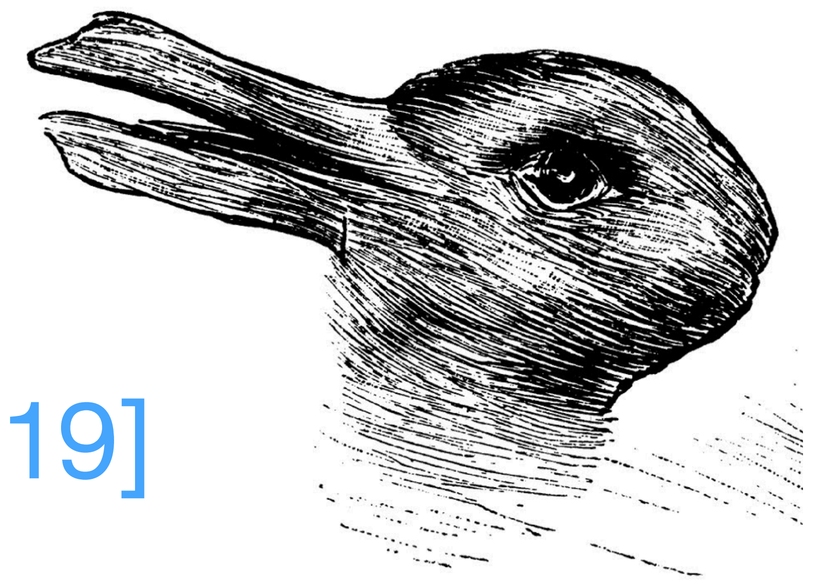
How can we generalize across diverse task structures and task goals?

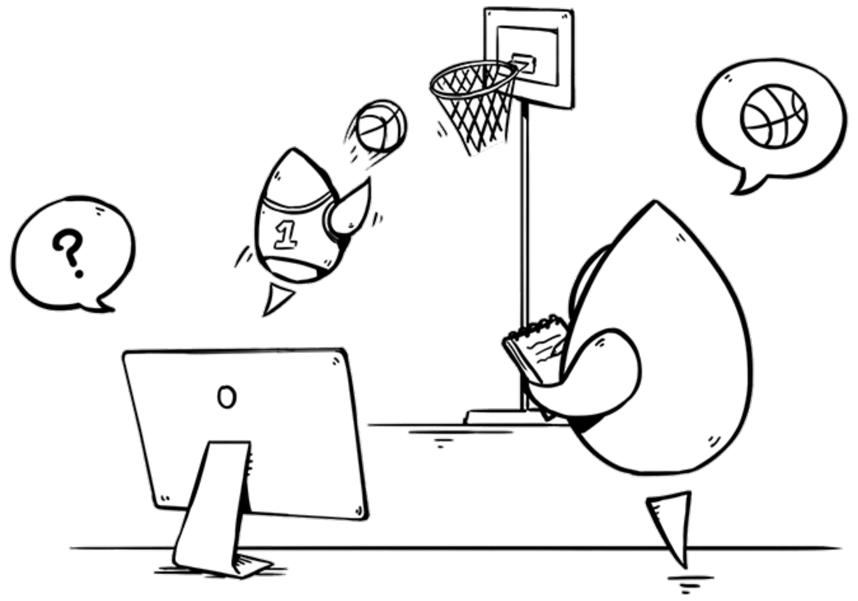
Using neural task graph as compositional inductive bias [CVPR'19]

Perceptual Uncertainty

How to address perceptual uncertainty arising from visual imitation?

Continuous relaxation of symbolic planner for one-shot imitation [IROS'19]





Long-horizon Tasks

How can we extrapolate to long-horizon tasks?

Symbolic regression planning with deep learning [NeurIPS'19]

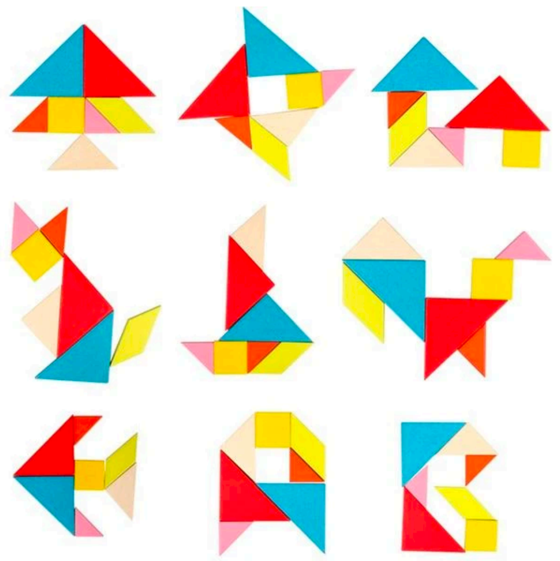

De-An Huang


Danfei Xu


Suraj Nair

Conclusions

Visual imitation Learning from video demonstrations



Compositional Generalization

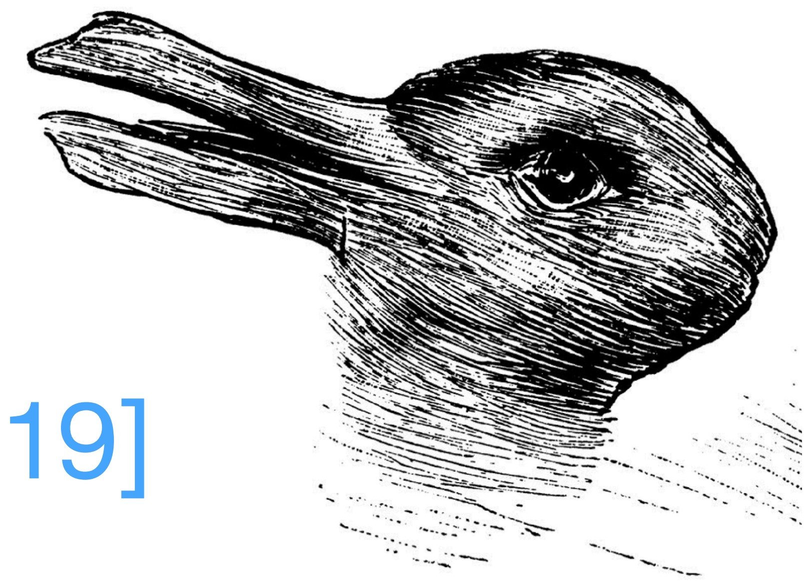
How can we generalize across diverse task structures and task goals?

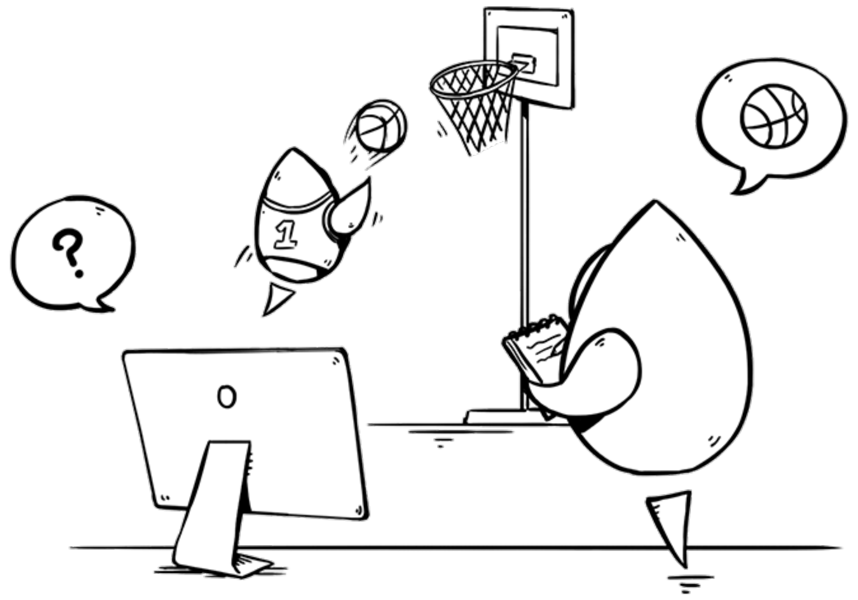
Using neural task graph as compositional inductive bias [CVPR'19]

Perceptual Uncertainty

How to address perceptual uncertainty arising from visual imitation?

Continuous relaxation of symbolic planner for one-shot imitation [IROS'19]





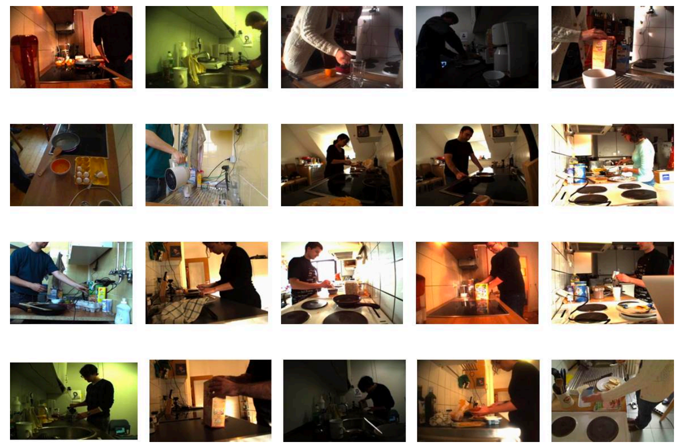
Long-horizon Tasks

How can we extrapolate to long-horizon tasks?

Symbolic regression planning with deep learning [NeurIPS'19]

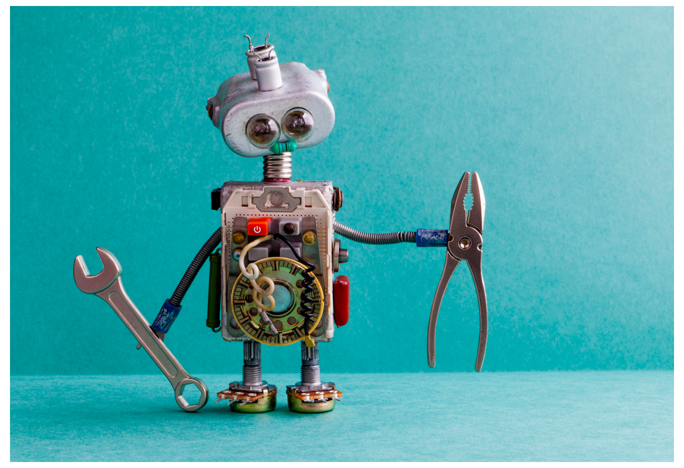
Open Challenges

Towards **web-scale visual imitation learning**



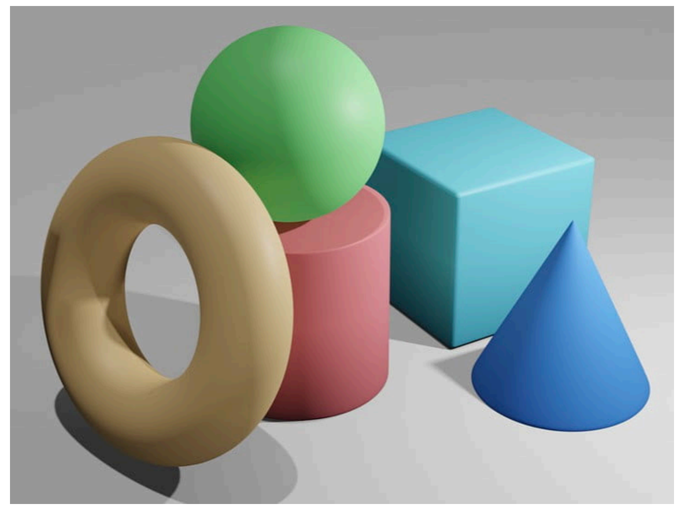
Imitation learning meets activity understanding.

Extracting meaningful task knowledge from unconstrainted web video data.



Where are the symbols coming from?

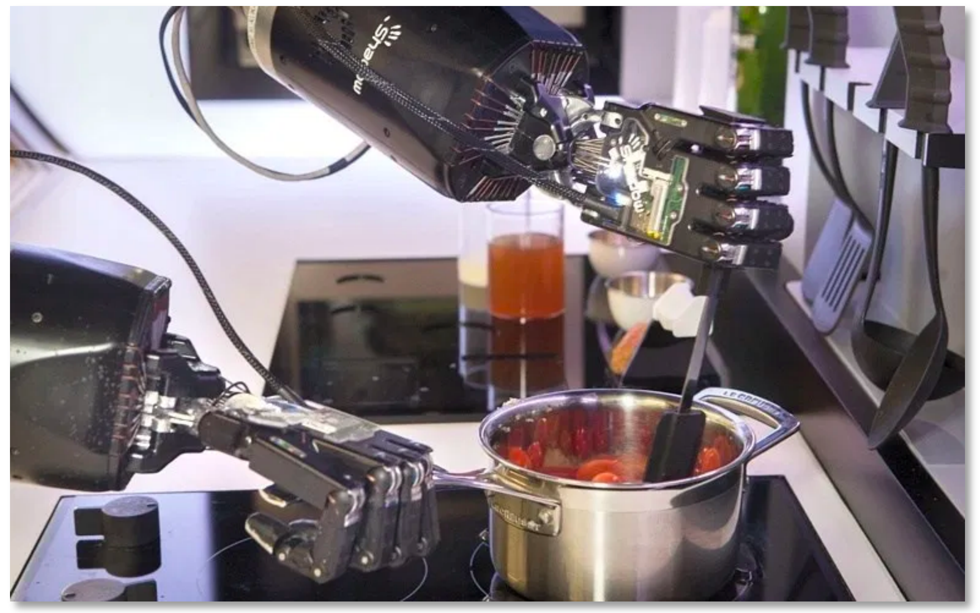
Concept learning and symbol discovery from self-supervised active exploration.



The next generation of hybrid AI systems

Imitation learning algorithms that seamlessly integrate neuro-symbolic hybrid methods.

video demonstration



robot execution