Learning for Planning and World Modeling: Structured Approaches

Linfeng Zhao

Northeastern University, Khoury College of Computer Sciences

http://lfzhao.com/

Brown Robotics Seminar, 2024/11


Hey Spot, put away the tools!
BostonDynamics

Why Planning?

- Difficulties in Perceiving and Representing a Complex World
- Acting in Complex Environments
- Generalization to Unseen Cases and Tasks
- Inference-Time Adaptivity

"They are one of a few groups that actually know planning and abstraction!"

Why Planning?

- Difficulties in Perceiving and Representing a Complex World
  - *Enabling operating with less perfect representation by considering more outcomes*
  - *Allowing making decisions to reduce uncertainty, partially observability and other imperfections in state*
- Acting in Complex Environments
- Generalization to Unseen Cases and Tasks
- Inference-Time Adaptivity

Why Planning?

- Difficulties in Perceiving and Representing a Complex World
- Acting in Complex Environments
    - *Strategic decisions for long-horizon and large action space*
    - *Handle various forms of goals and subgoals and strategic exploration*
- Generalization to Unseen Cases and Tasks
- Inference-Time Adaptivity

Why Planning?

- Difficulties in Perceiving and Representing a Complex World
- Acting in Complex Environments
- Generalization to Unseen Cases and Tasks
  - *Enabling generalization to unseen situations and novel combinations of known skills*
  - *Allowing agents to adapt to new tasks and environments*
- Inference-Time Adaptivity

Why Planning?

- Difficulties in Perceiving and Representing a Complex World
- Acting in Complex Environments
- Generalization to Unseen Cases and Tasks
- Inference-Time Adaptivity
    - *Allowing for inference-time adaptivity, enabling agents to adjust their strategies on the fly and allocate computational resources efficiently during decision-making*

Why Learning for Planning?

Pure Learning

- A.k.a. using a big neural network trained with gradient descent on tons of data!
- It is too hard to learn with limited robot-specific data.

Pure Planning

- It involves numerous engineered components specific to tasks.
- It is hard to directly apply to open-world settings that robot doesn't know ahead of.

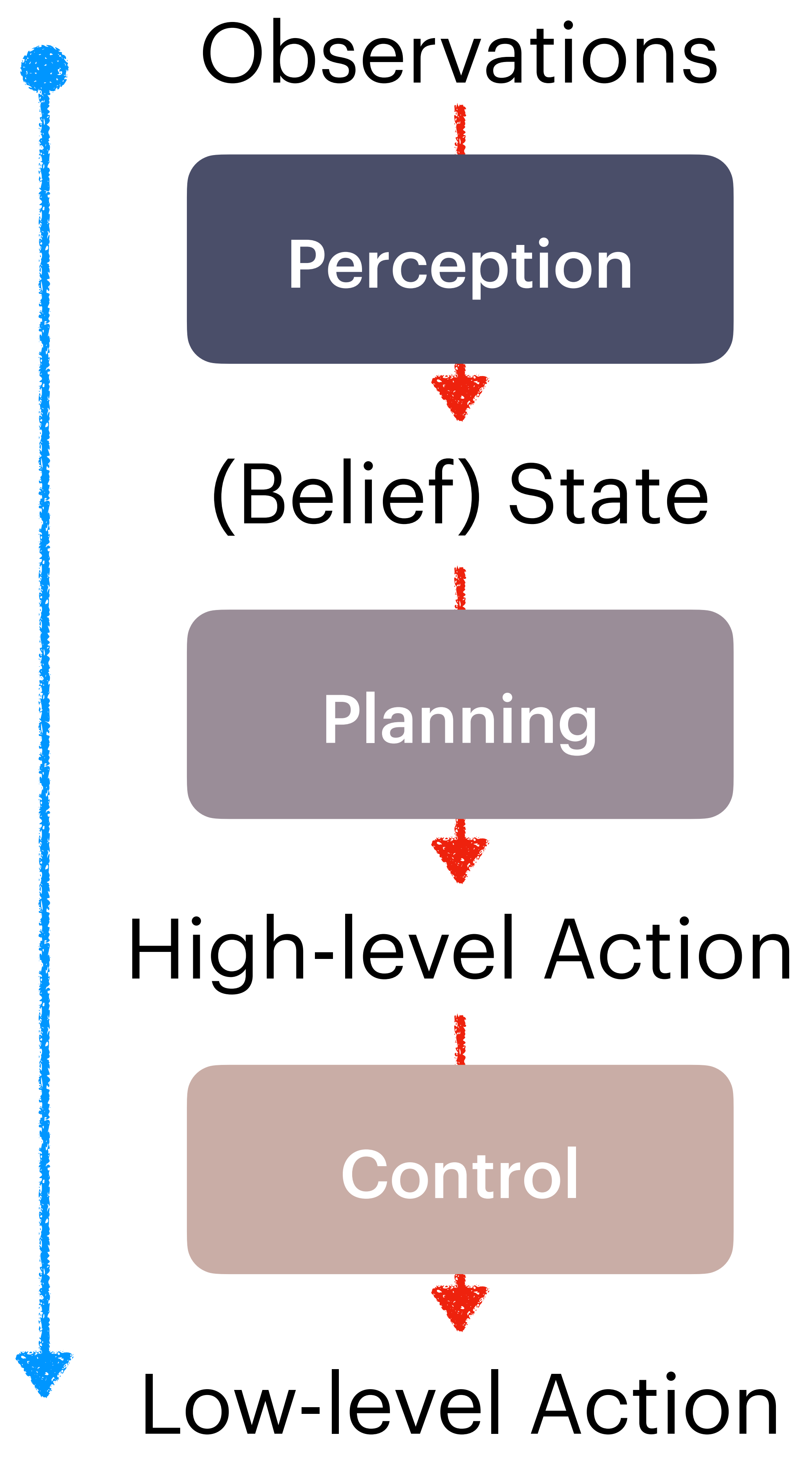
Can we keep the flexibility of learning while the benefits of planning?

Outline

- **Background**: Learning for Planning and World Modeling
- **Part 1**: ***Lossless*** Abstraction of World Representation and Planning
    - *Typically within a “flat” MDP*
    - E.g,. Geometric Structure, including symmetry and compositionality
- **Part 2**: ***Lossy*** Abstraction of World Representation and Planning
    - *Typically with hierarchies, such as high-level and low-level*
    - E.g., Using symbols/language or maps

Background:
Learning for Planning and World Modeling

Modeling the Environment



Typically modeled as a (PO)MDP

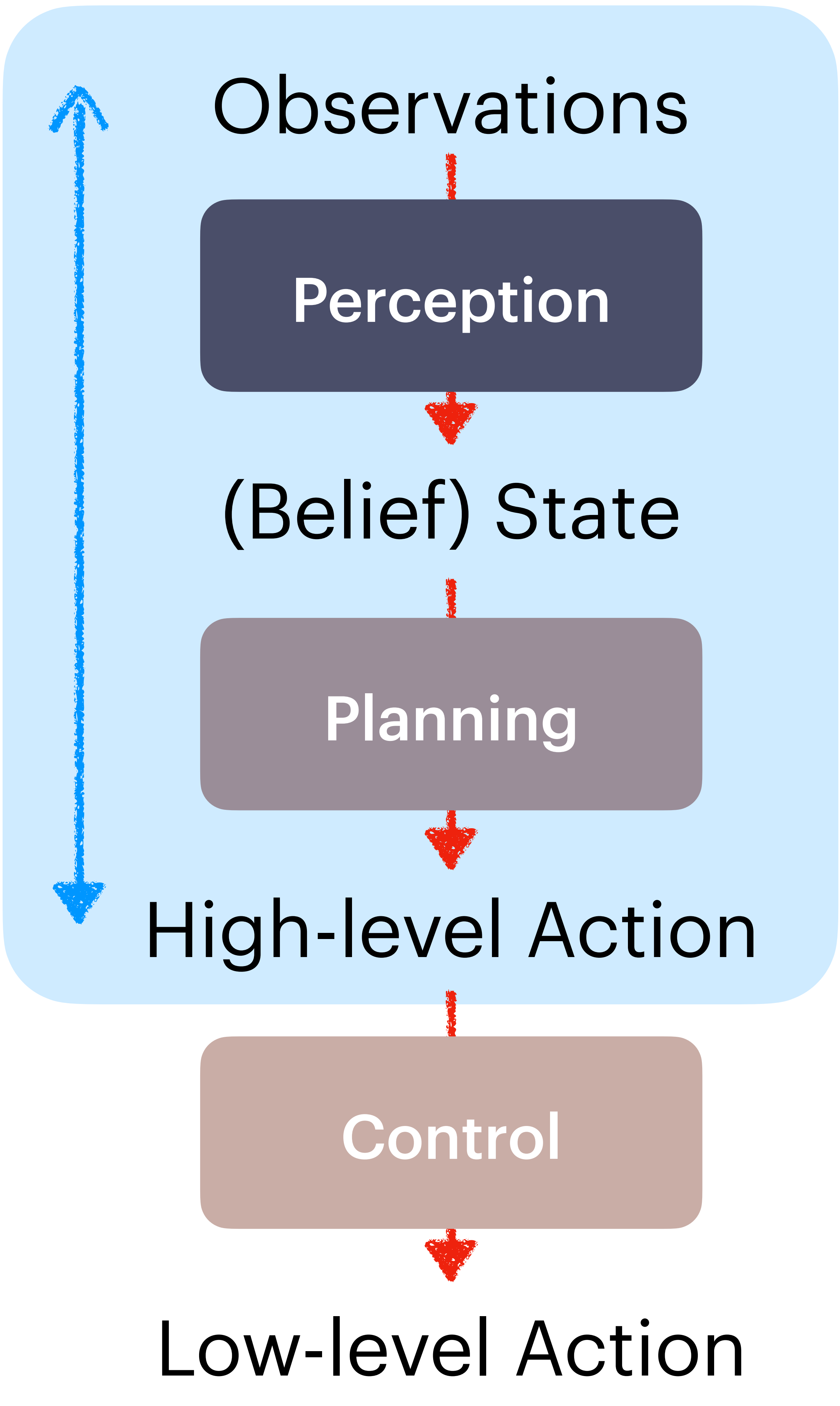
A classic robotics stack:

- Perception: *Extract features from past observations*
- Planning: *Produce sequence of high-level actions*
- Control: *Ground into low-level motor actions*

A naive approach: End-to-end learning $a_t = \pi_\theta(s_t)$

Inefficient, Poor Generalization, and Brittle

Modeling the Environment



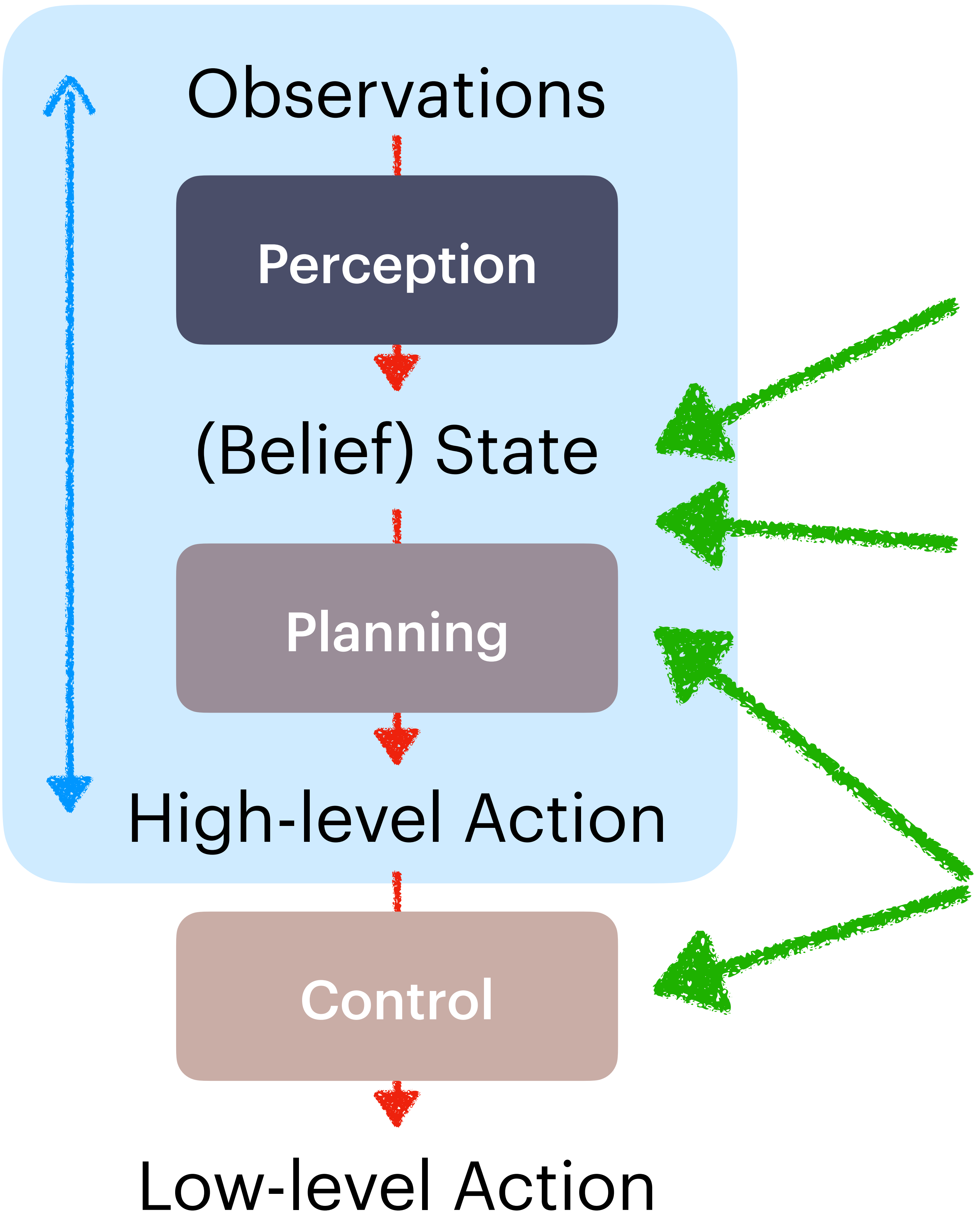
Typically modeled as a (PO)MDP

A classic robotics stack:

- Perception: *Extract features from past observations*
- Planning: *Produce sequence of (high-level) actions*
- Control: *Ground into low-level motor actions*

My research:
Using Learning for World Modeling and Planning with Proper Structure

World Modeling and Planning



World Modeling:

Building good state representation $s_t = h(o_{1:t})$

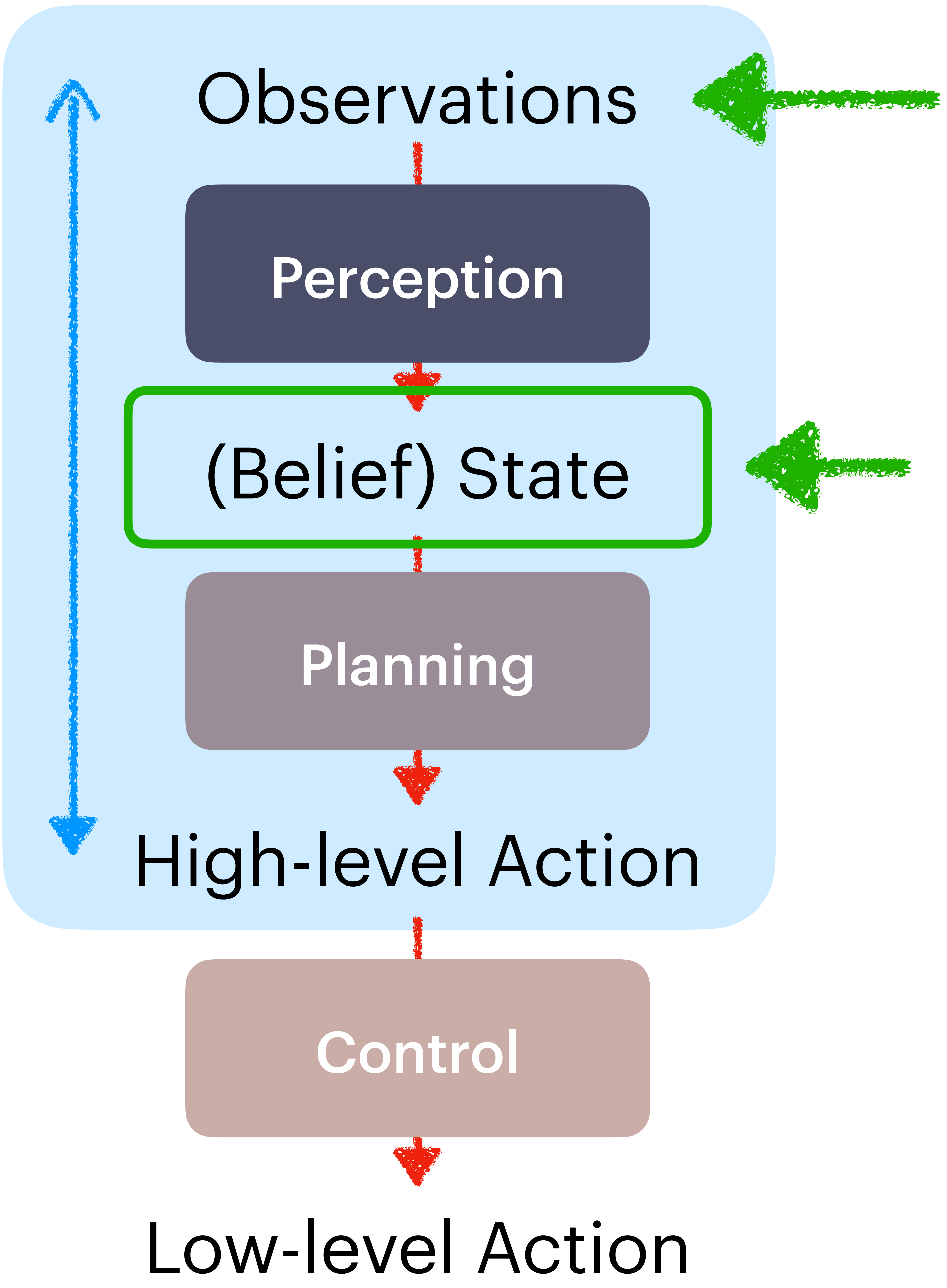
(e.g,. learning a vision encoder / gathering information)

Building good dynamics model $s_{t+1} = f(s_t, a_t)$

Planning:

Producing sequence of actions $a_{t:t+H} = \pi(s_t, g)$

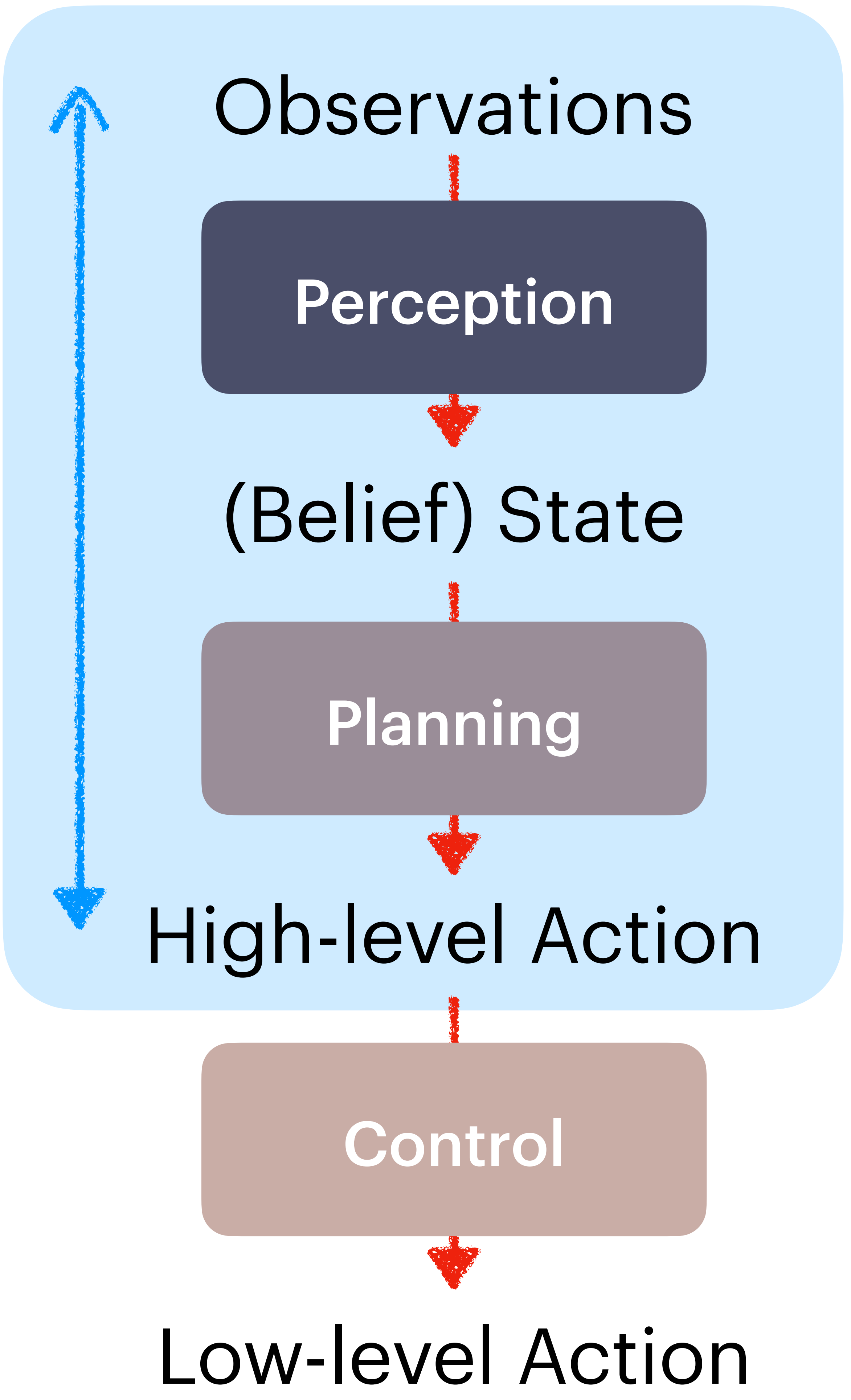
Why Joint World Modeling and Planning?

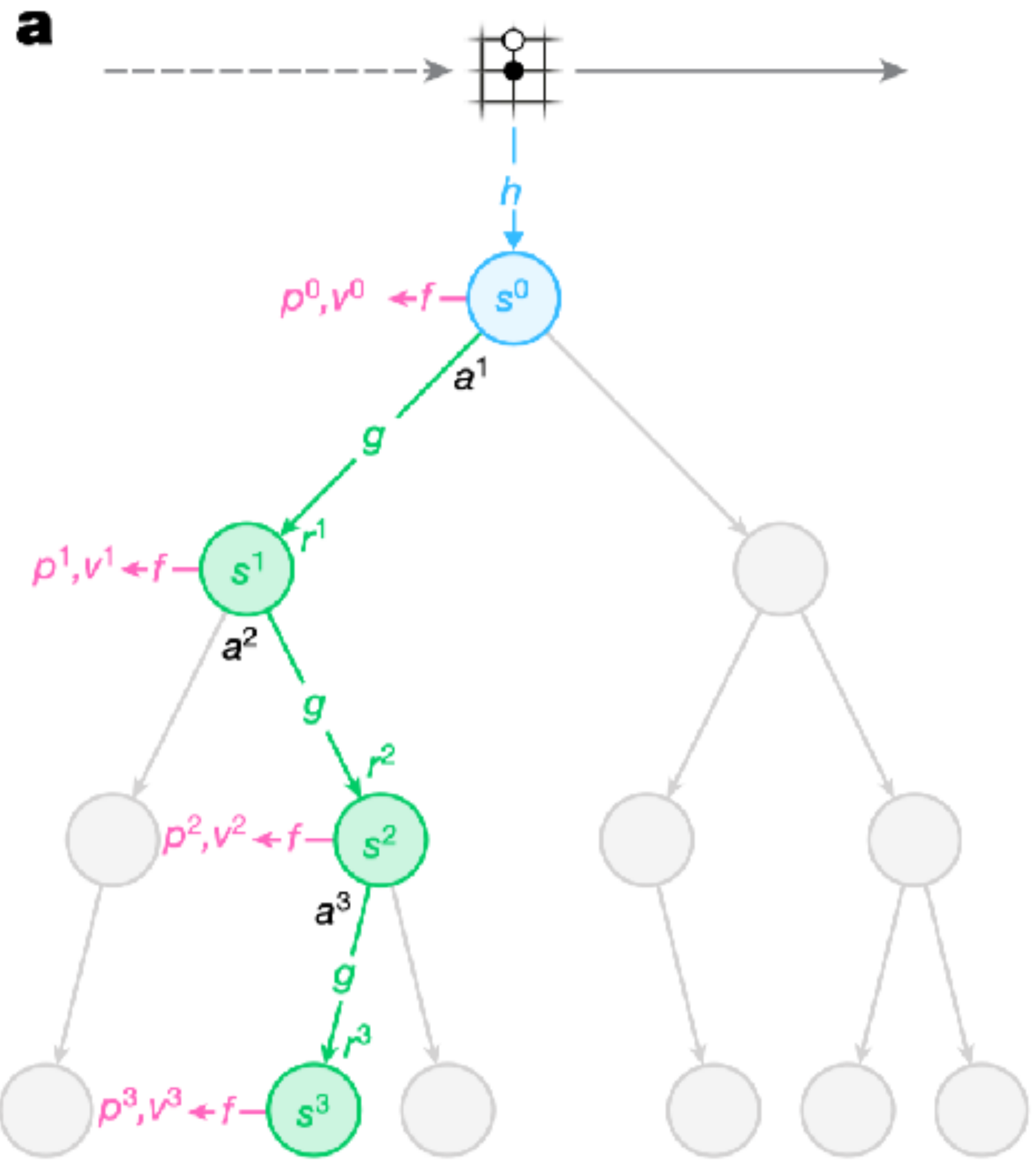


Further Challenge:
It is challenging to pre-specify state representation for open-world (truly unseen) environments.

Example:
A robot tasked to remove some tools in a cabinet.
It needs to represent uncertainty in object existence and properties.

How Joint World Modeling and Planning?

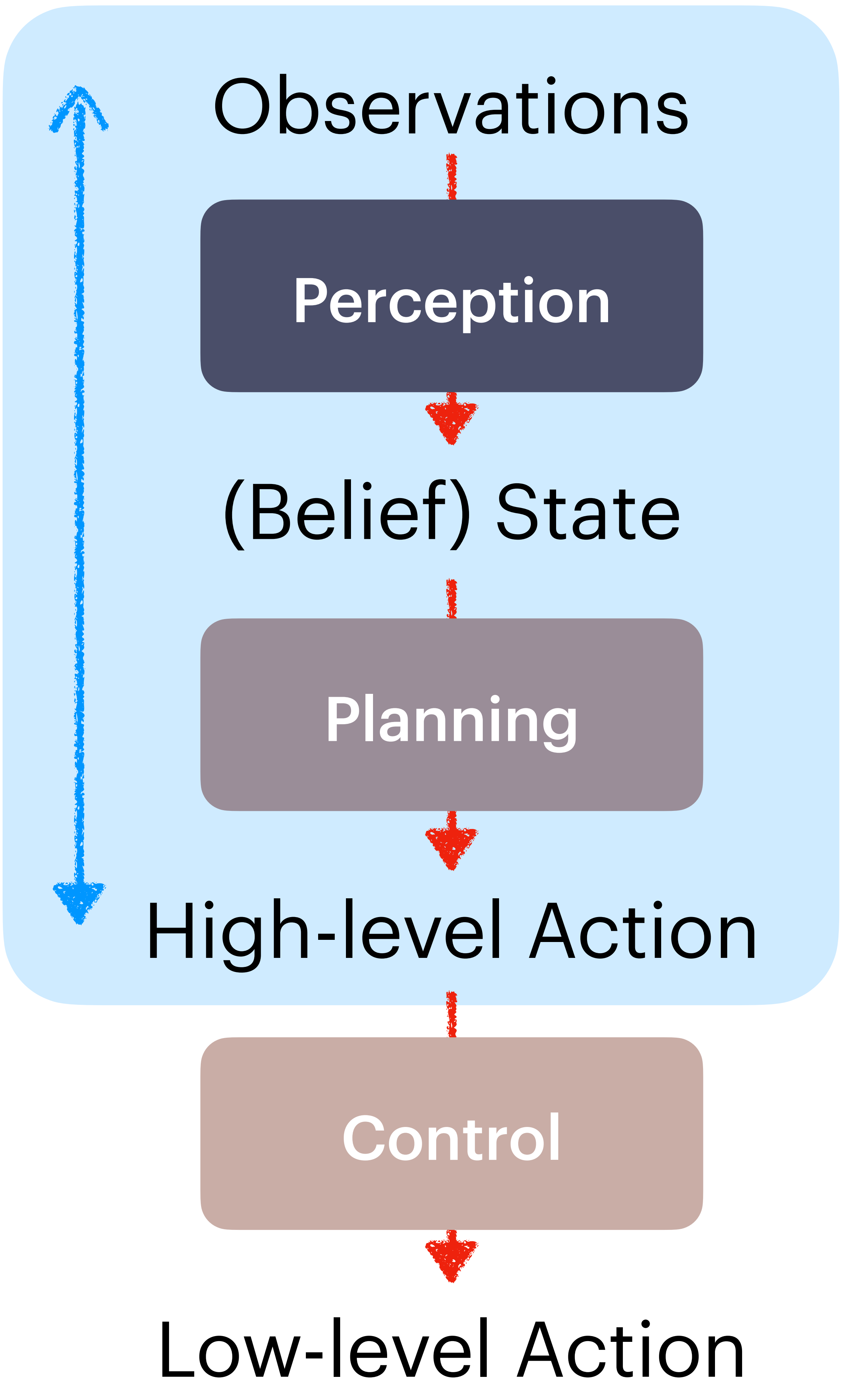




Previous: MuZero for Go (Schrittwieser et al. 2019), ...

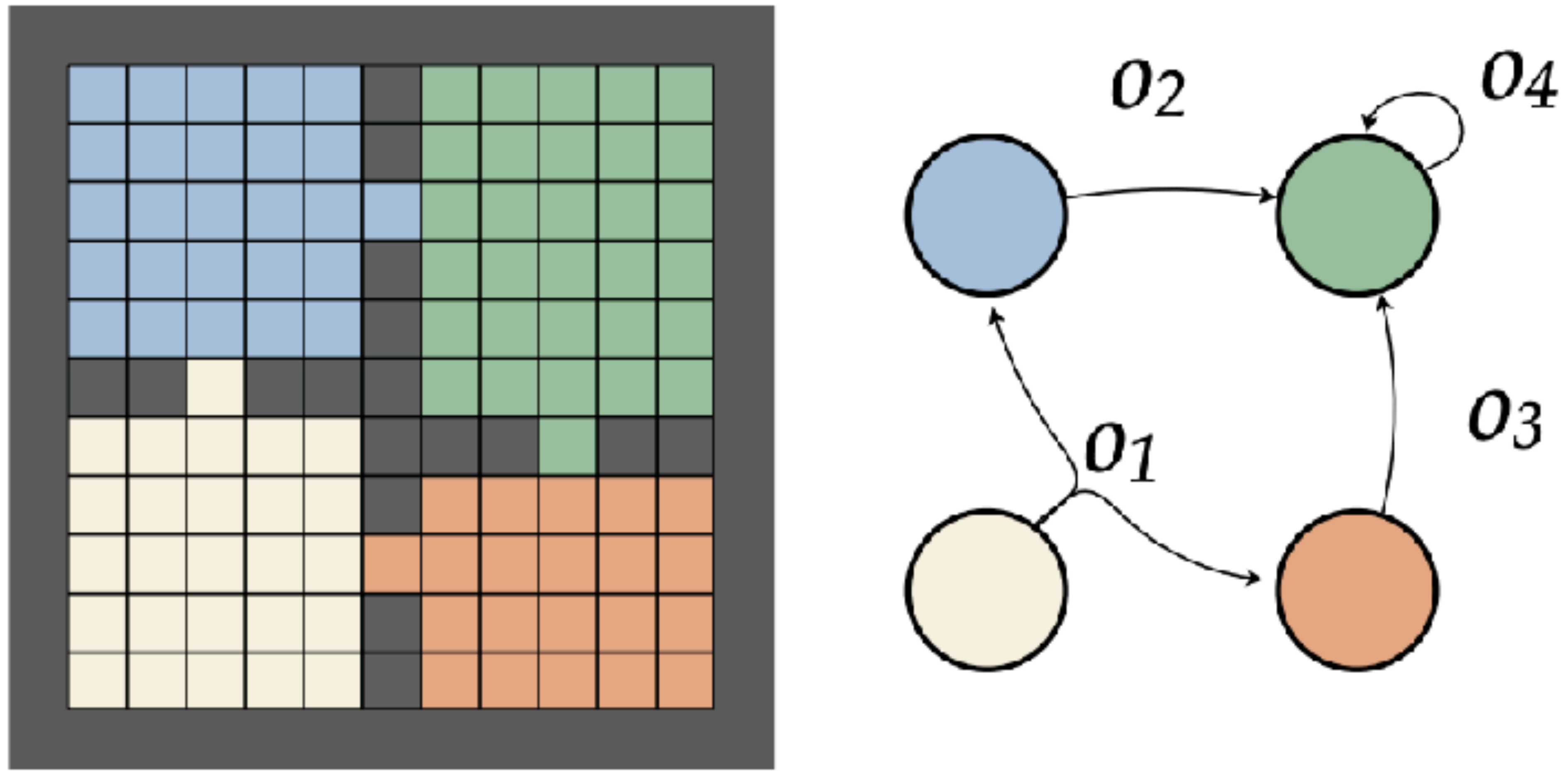
Recent: VLM for task planning (E.g., Driess et al. 2023), ...

Challenges and Considerations

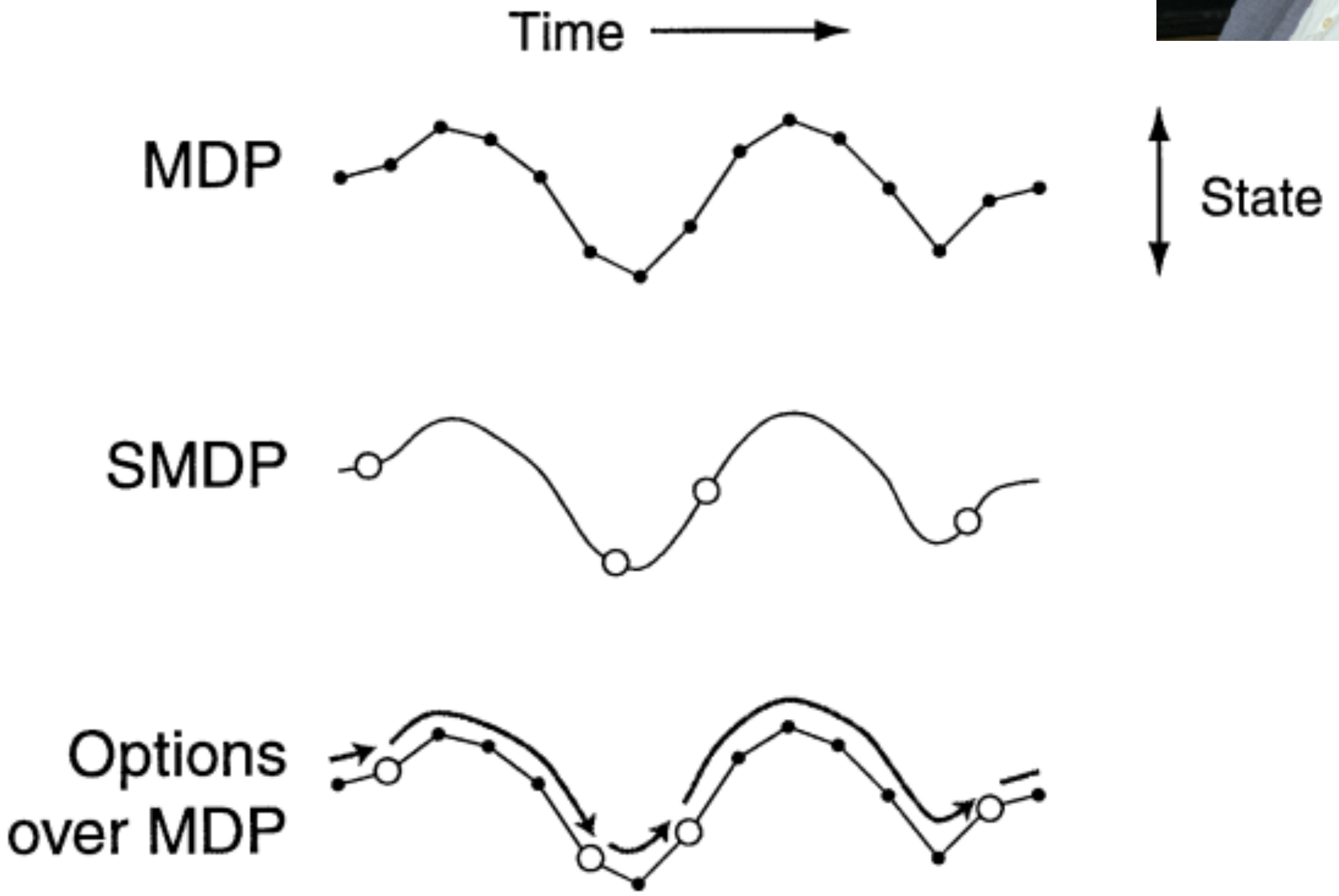


- Promising; but needs lots of data to train!
- Can we make use of the structure of the tasks and the algorithmic stack to reduce the complexity?
- ***More Structure in Algorithms***
  - Consider the training data sources and model complexity for different components
- ***More Structure in Tasks, e.g., abstraction***
  - Exploit efficient structure from environments and tasks to

State & Action Abstraction



(a) Assignment of options to each s_ϕ via $\pi_{\mathcal{O}_\phi}$.

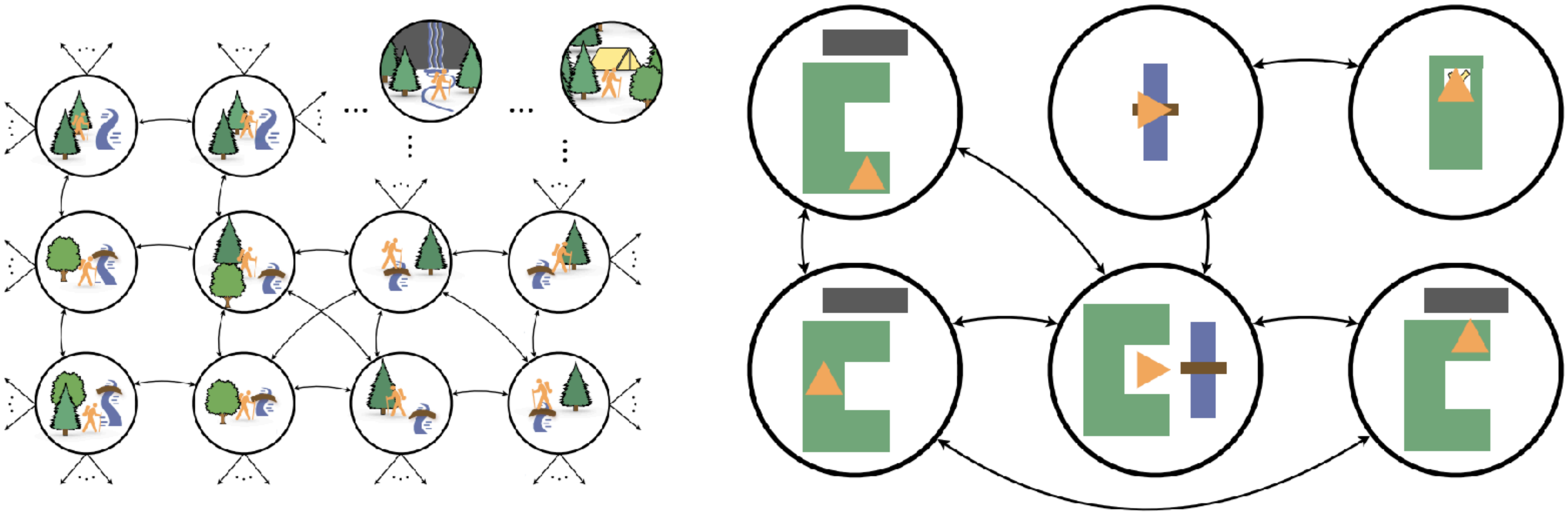


State Abstraction
(State Partition)

Action Abstraction
(Options)

Reinforcement Learning: An Introduction. Andrew Barto and Richard S. Sutton 2018.
A Theory of Abstraction in Reinforcement Learning. David Abel, PhD Thesis 2020.

Planning in the abstracted model



(a) Reasoning in the environment. (b) Reasoning in the abstract.

A Theory of Abstraction in Reinforcement Learning. David Abel, PhD Thesis 2020.

Lossless vs. Lossy Abstraction

State Abstraction as Compression in Apprenticeship Learning

David Abel,[1] Dilip Arumugam,[2] Kavosh Asadi,[1]
Yuu Jinnai,[1] Michael L. Littman,[1] Lawson L.S. Wong[3]
[1]Department of Computer Science, Brown University
[2]Department of Computer Science, Stanford University
[3]College of Computer and Information Science, Northeastern University

David Abel

Dilip Arumugam

Kavosh Asadi

Yuu Jinnai

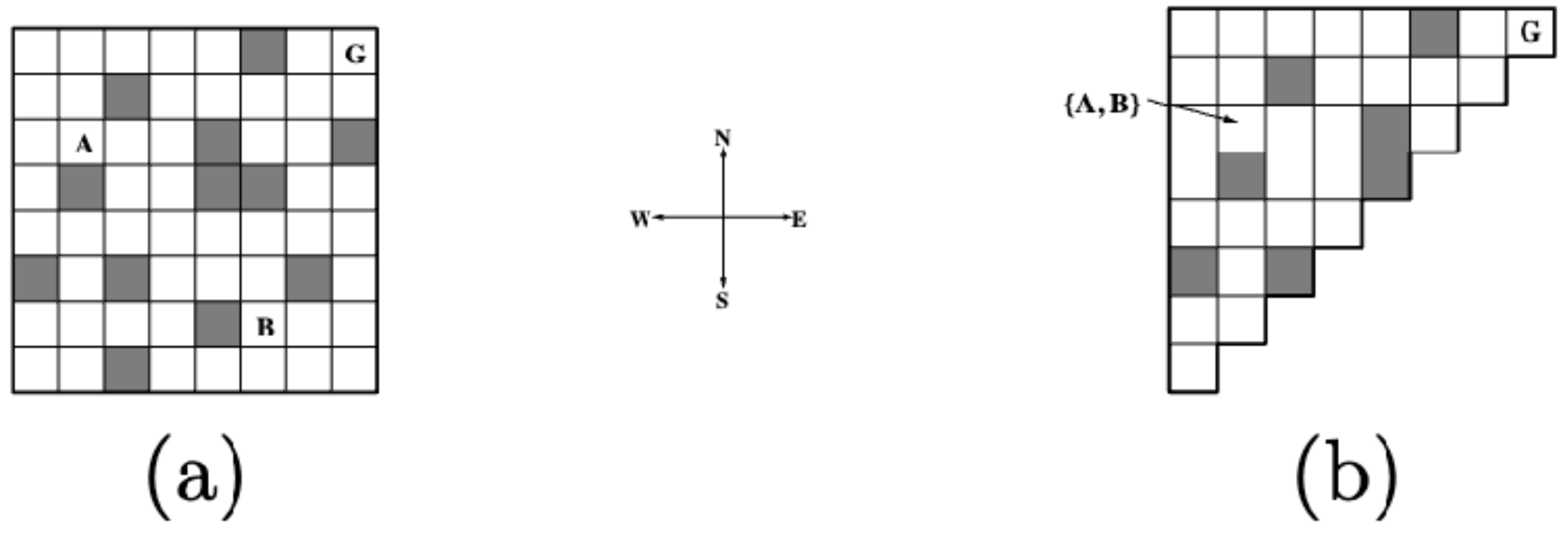
Michael Littman

- View state representation using terminology in ***compression***
- It decides the behaviors of planning:
  - ***Lossy*** representation may need additional ***grounding*** of abstract plan to the original space

Part 1: *Lossless* Abstraction of World Representation and Planning

Lossless Abstraction
MDP Homomorphism and Value Equivalence

- "Lossless" here refers to abstraction that preserves optimal ground policy
- *MDP Homomorphism* defines a mapping from the ground MDP to a reduced MDP that preserves its "structure"
  - Including optimal values and policies



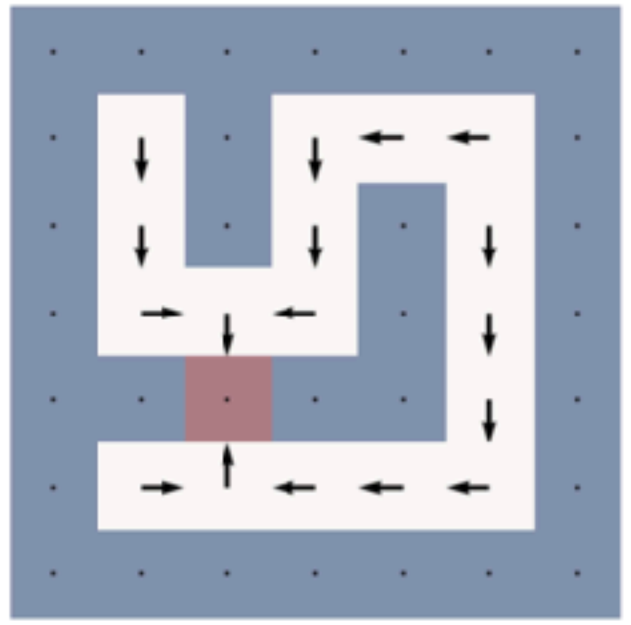
A symmetric gridworld problem.

"A" and "B" states are equivalent because each action at A exists an equivalent action at B.

[Ravindran & Barto, An Algebraic Approach to Abstraction in Reinforcement Learning, 2003]

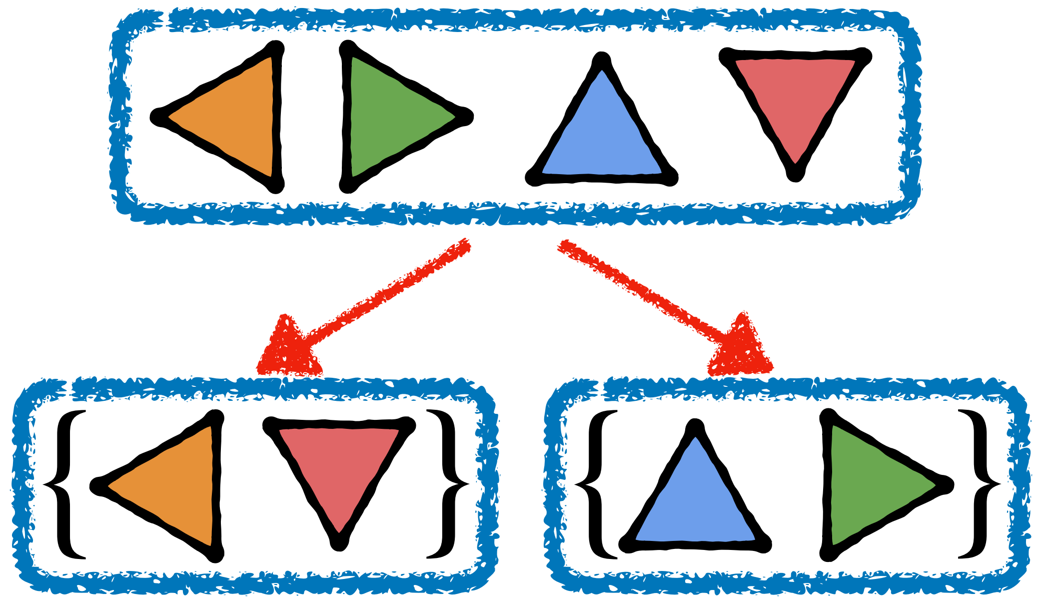
Lossless Abstraction Induced by “Structure”

- Use e.g,. symmetry and compositionality properties of the tasks to induce abstracted MDP
- May reduce number of free parameters and solution space
- Result in better efficiency, generalization, scalability, …



2D Discrete Map Rotation

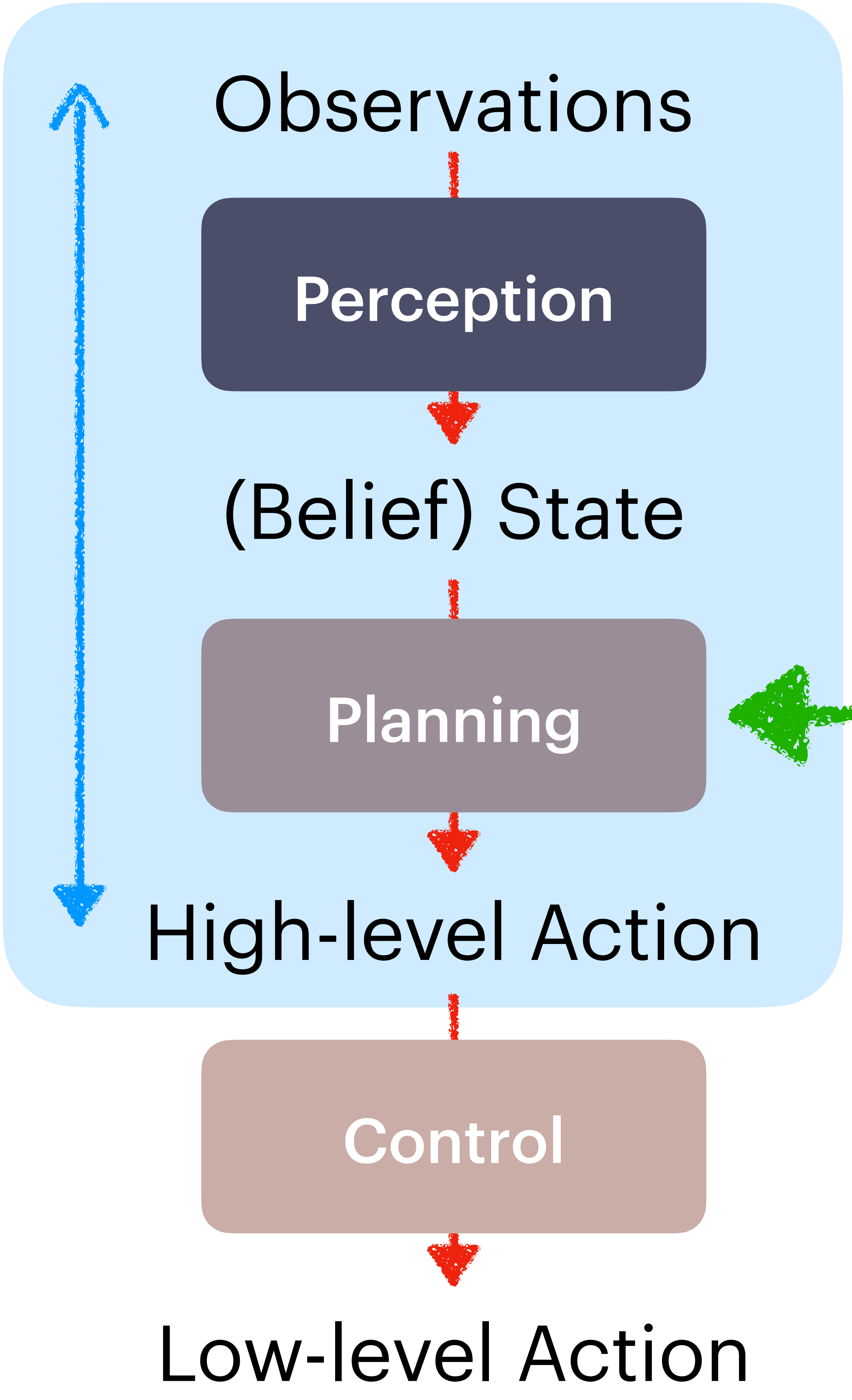
2D Discrete Symmetry D_4



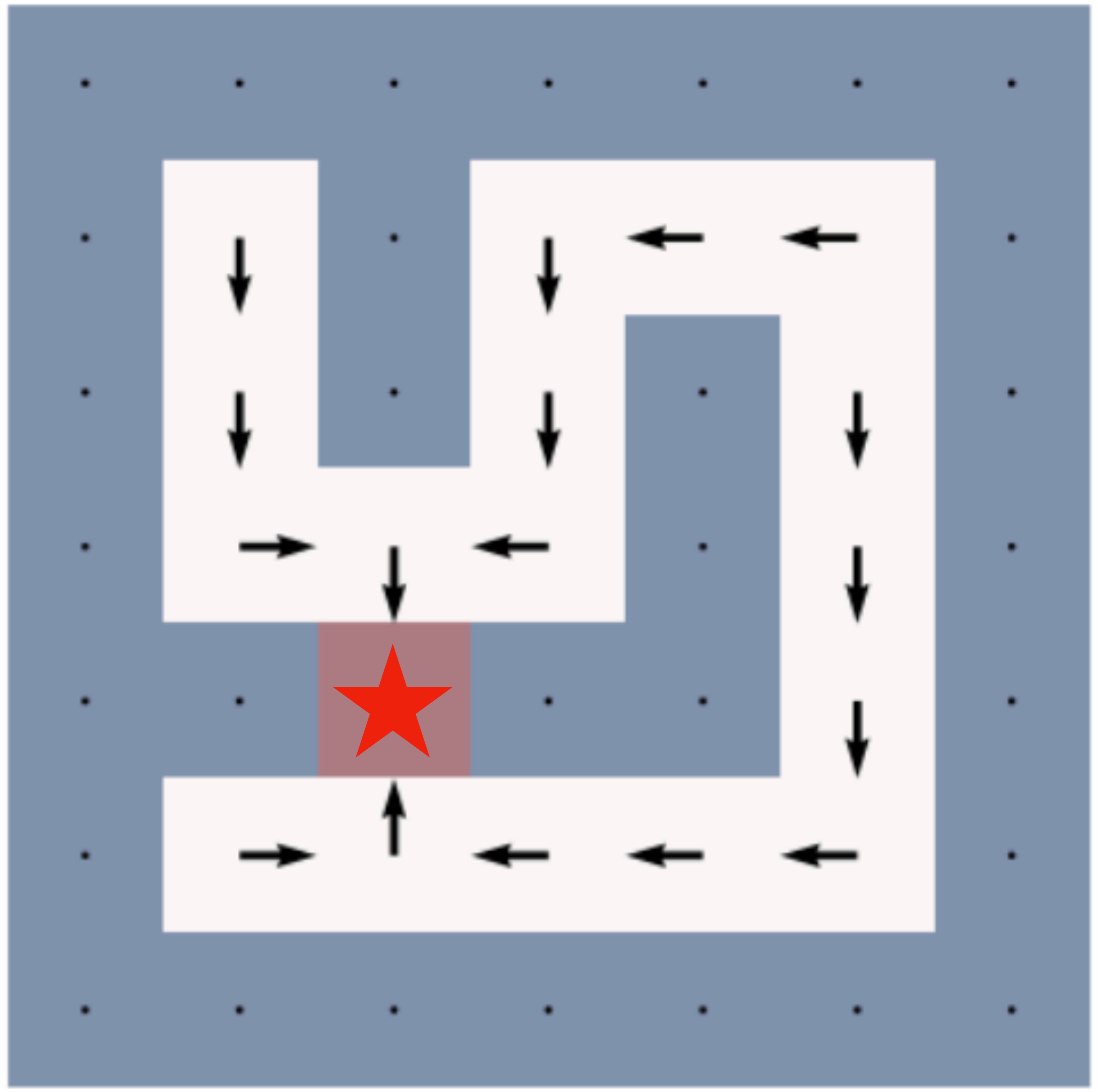
Object Interchangeability

Permutation Symmetry S_N

Integrating Symmetry into Planning

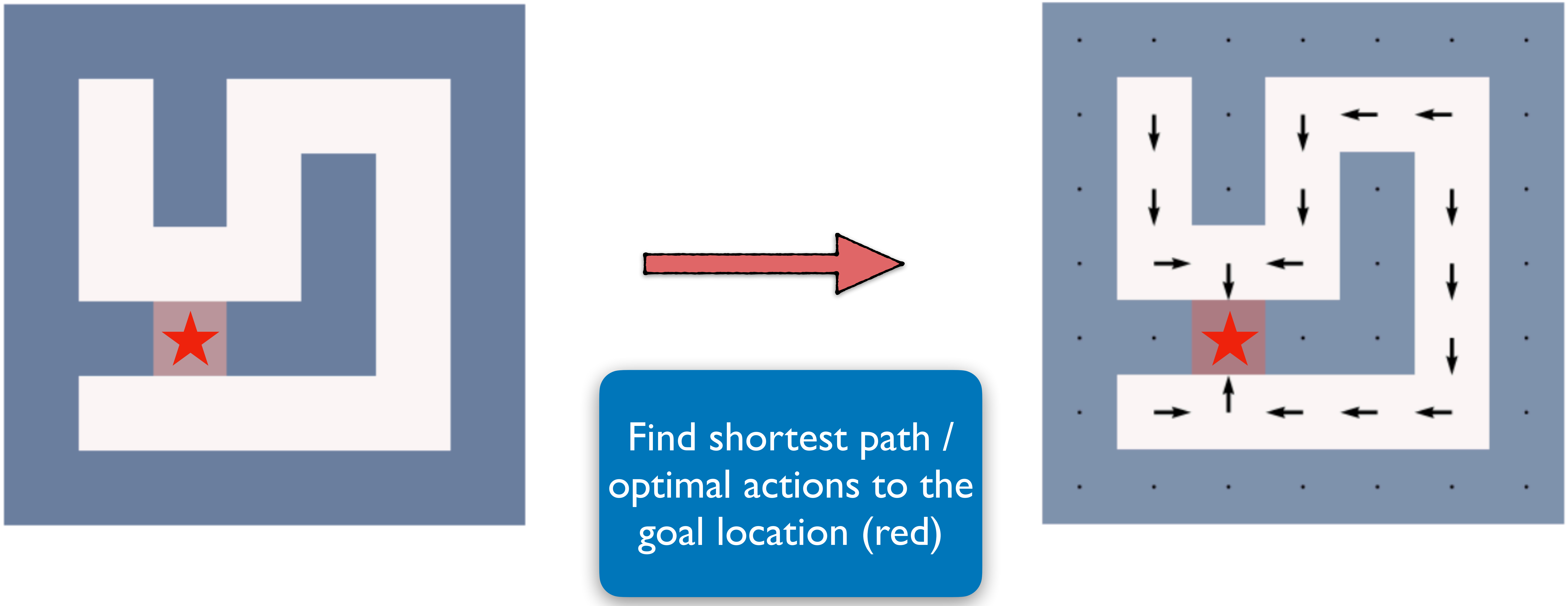


Can we use symmetry to improve learning-based planning, while keeping end-to-end learnable?



Zhao, Zhu, Kong, Walters, Wong. “Integrating Symmetry Into Differentiable Planning With Steerable Convolutions”. ICLR 2023.

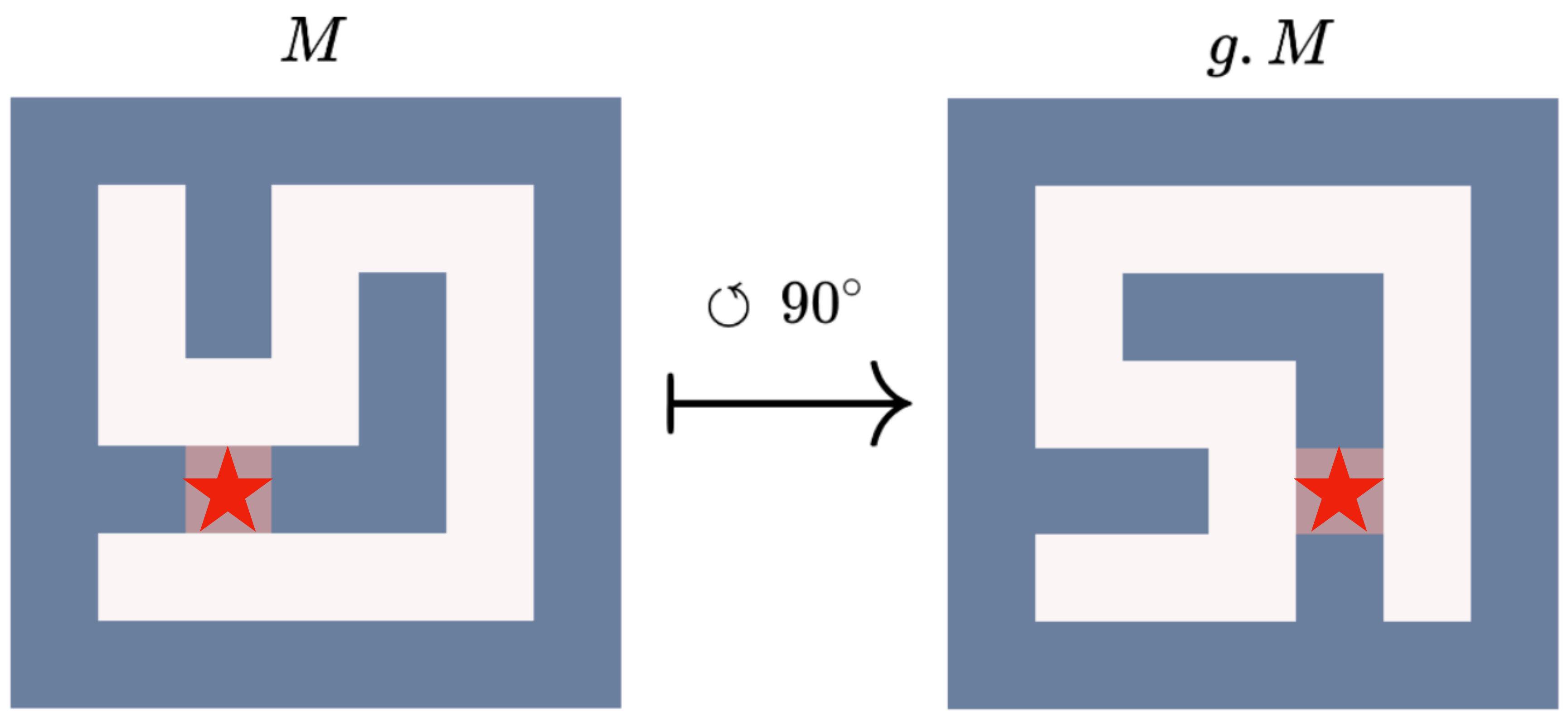
Path Planning



Zhao, Zhu, Kong, Walters, Wong. "Integrating Symmetry Into Differentiable Planning With Steerable Convolutions". ICLR 2023.

Symmetry in Path Planning

What does the symmetry look like?



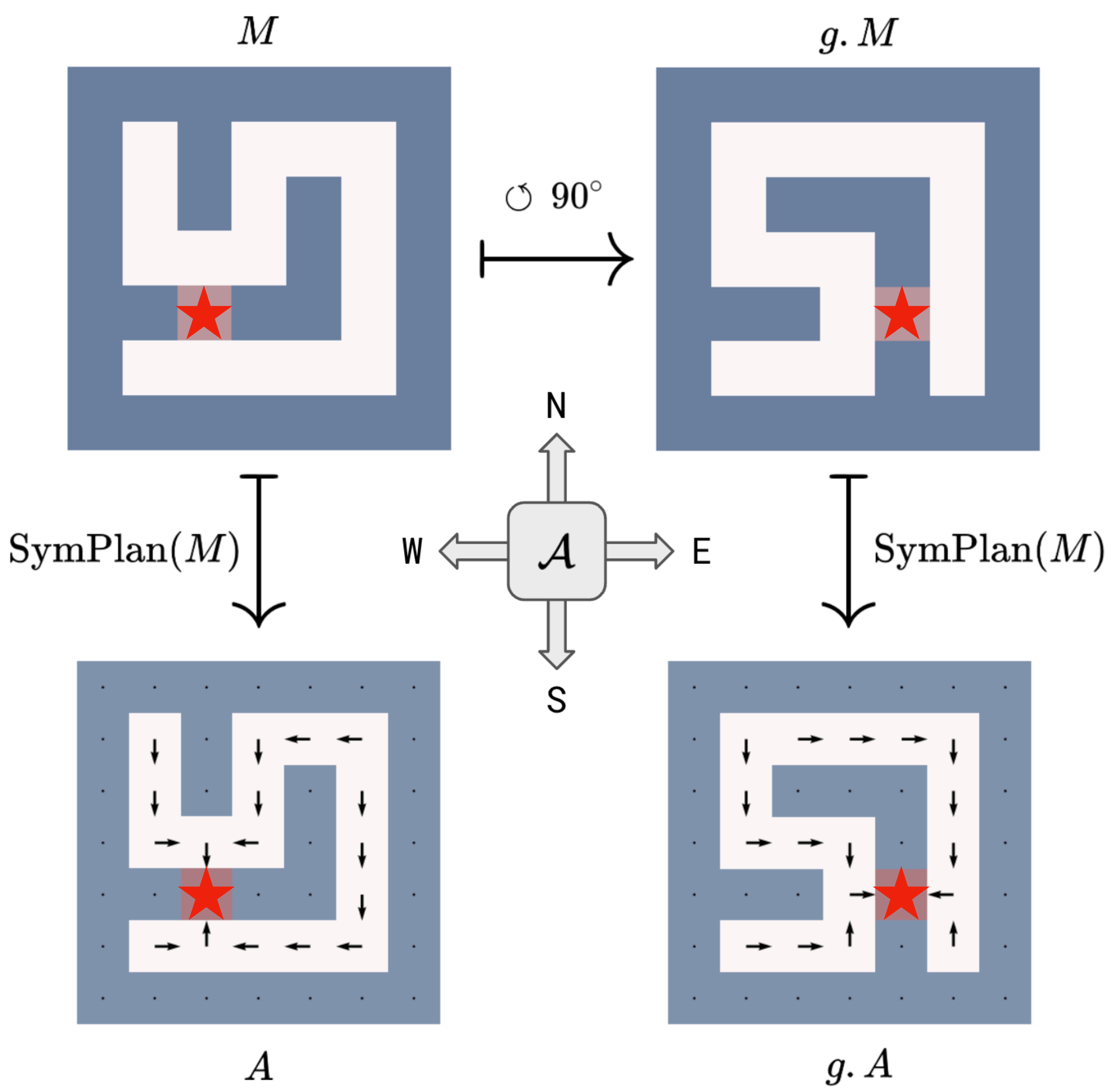
Symmetry in Path Planning

What does the symmetry look like?

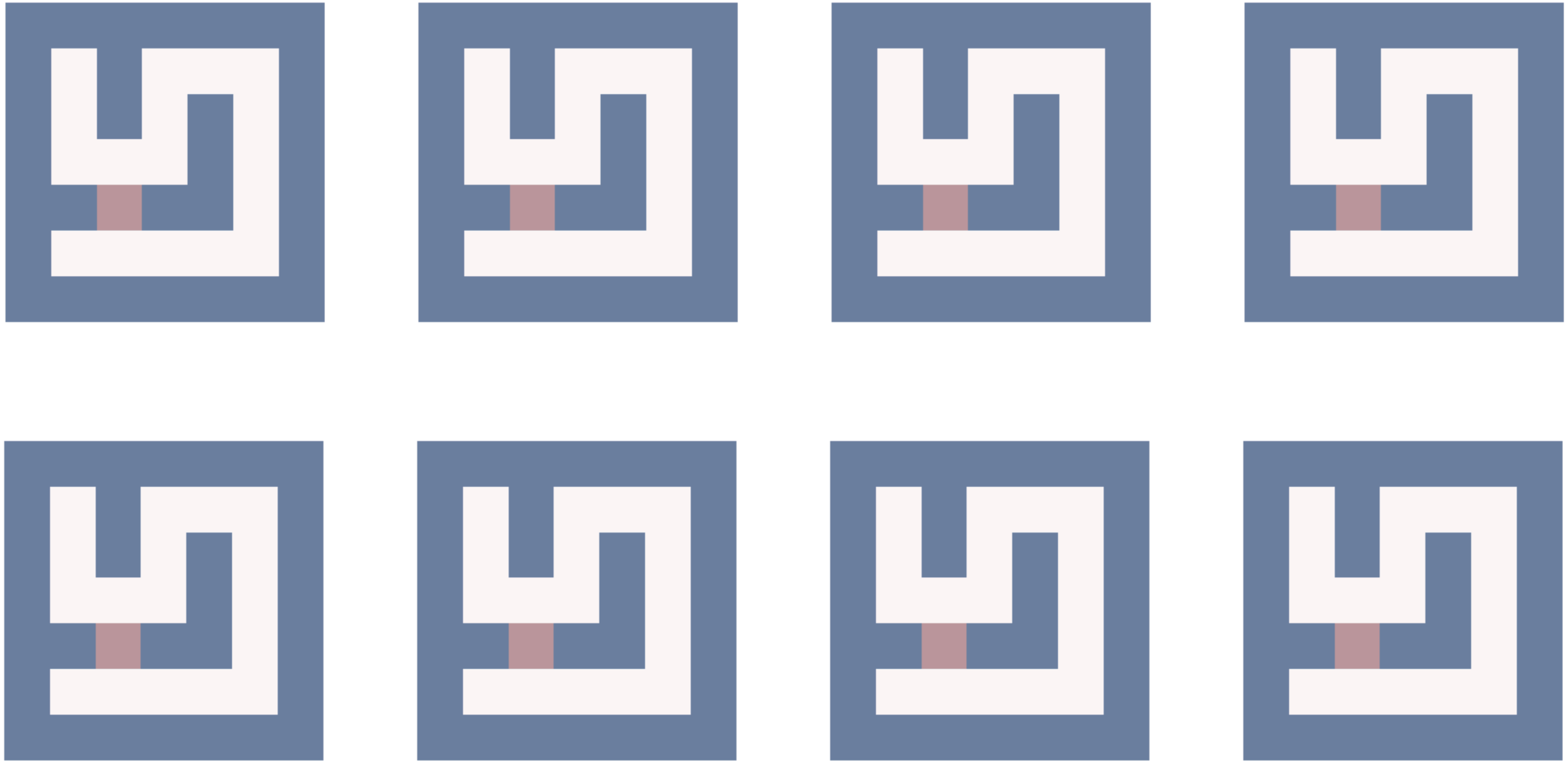
They can be described by

Equivariance

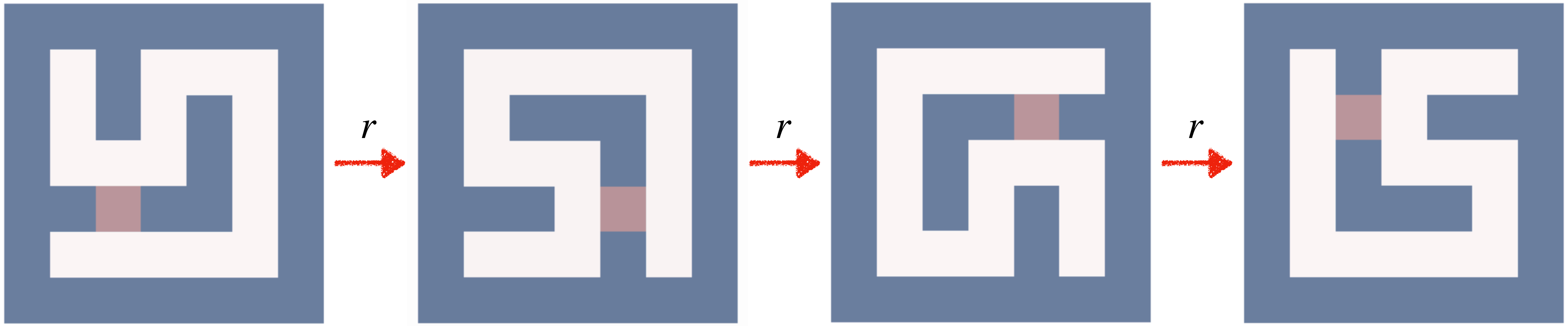
$$\circlearrowleft 90^\circ \circ \left(\text{Plan}(M)\right) = \text{Plan}(\circlearrowleft 90^\circ \circ M)$$



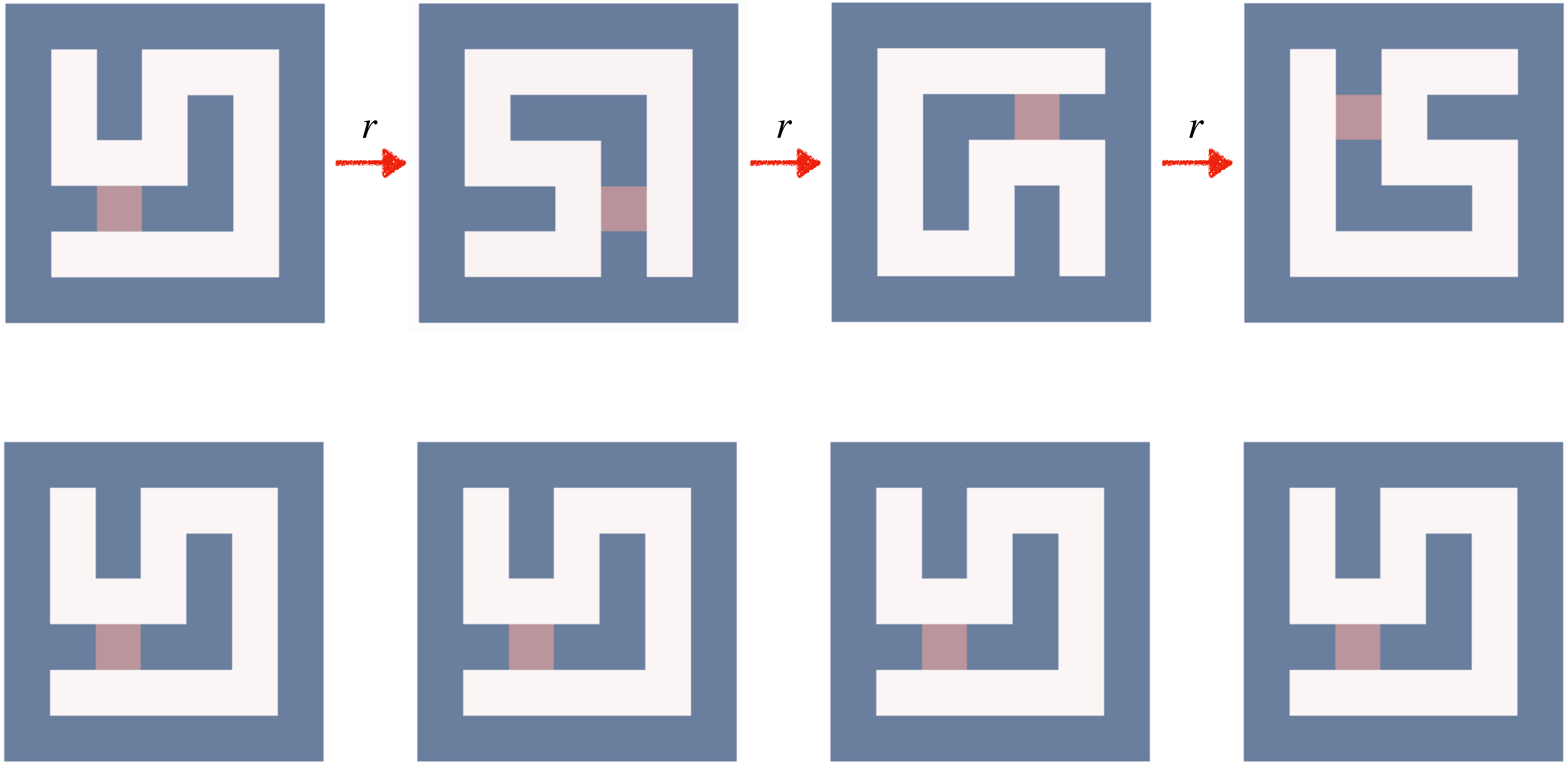
Symmetry: All Rotations and Reflections



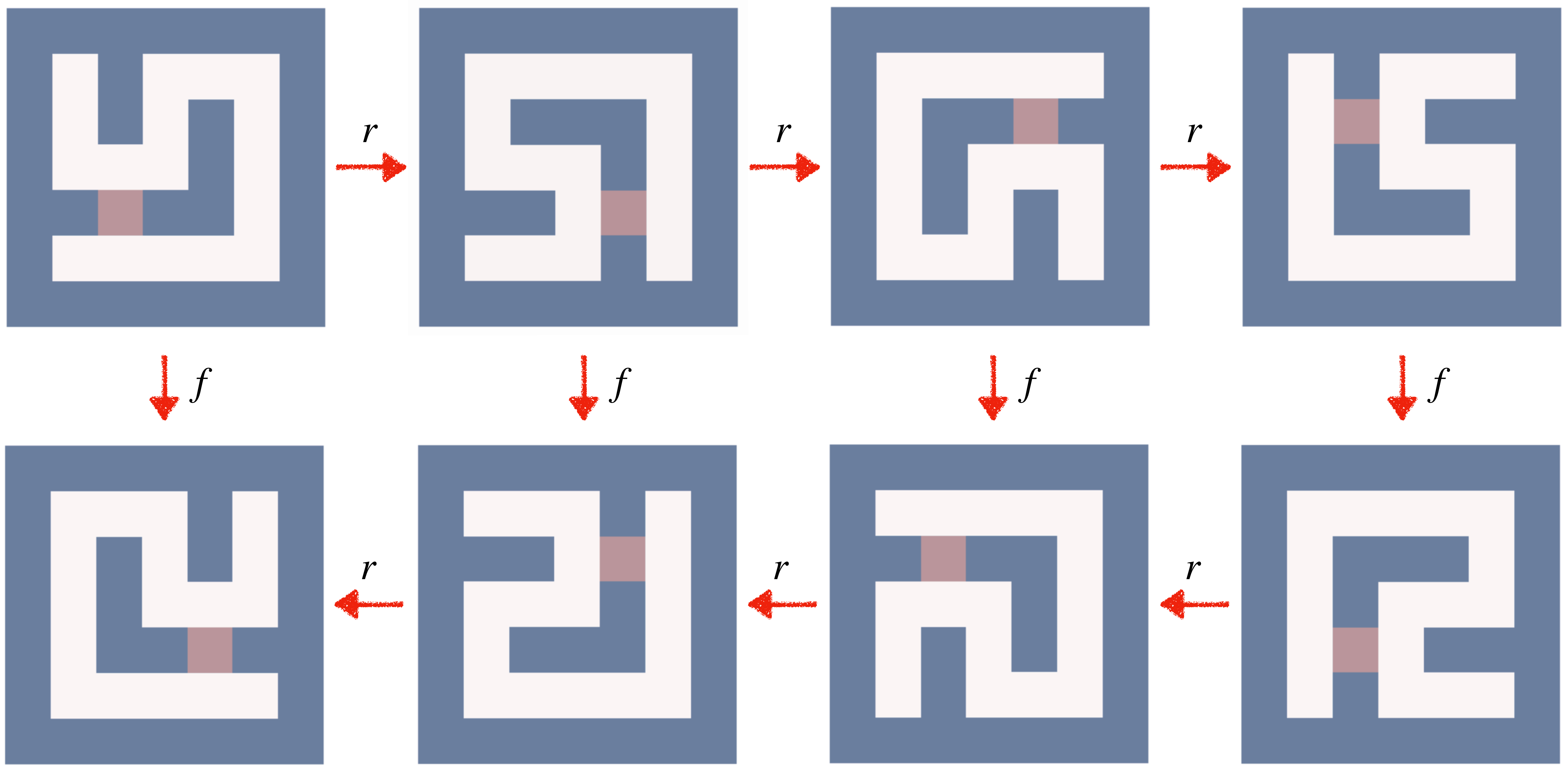
Symmetry: Rotations



Symmetry: Rotations and Reflections

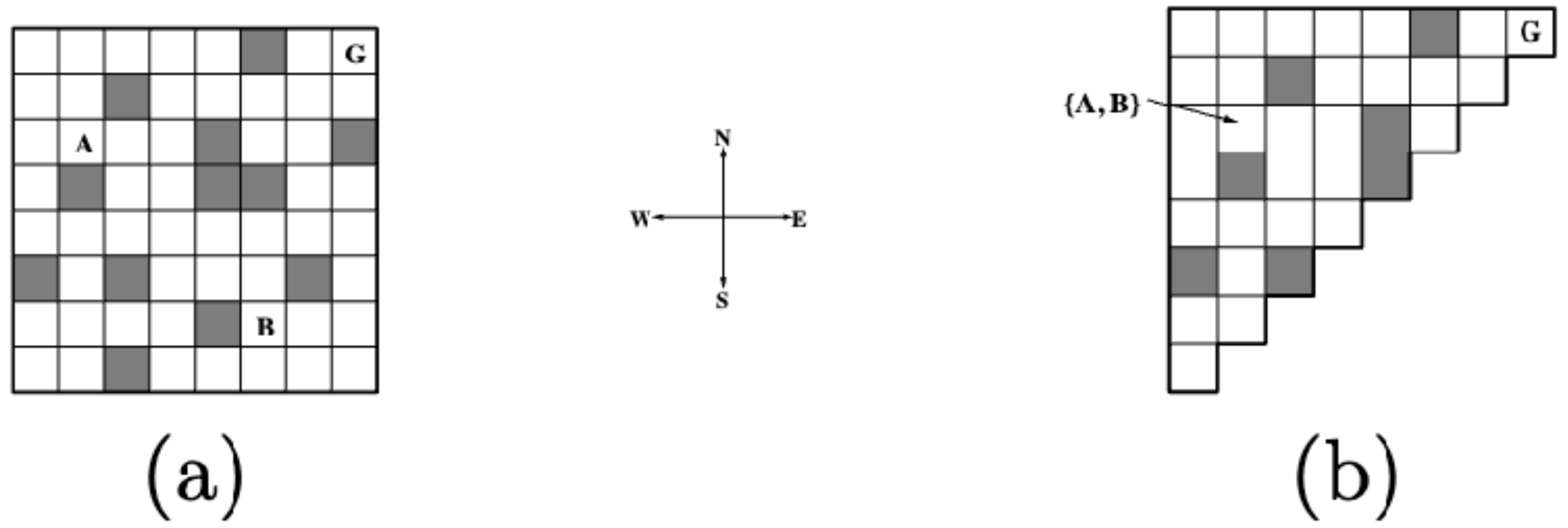


Symmetry: All 8 Transformations in D_4



Difficulties in Exploiting Symmetry

- Classical planning algorithms need explicit representation of the MDP $\overline{\mathscr{M}}$
    - E.g., transition dynamics $\overline{P}$
- Finding a homomorphic MDP is NP-hard problem
- Can we avoid:
  (1) NP-hard orbit search of equivalent state-action pairs and (2) explicitly representing the reduced MDP?



A symmetric gridworld problem.

"A" and "B" states are equivalent because each action at A exists an equivalent action at B.

[Ravindran & Barto, An Algebraic Approach to Abstraction in Reinforcement Learning, 2003]

Value Iteration with Symmetry

Every update is equivariant — Local Equivariance

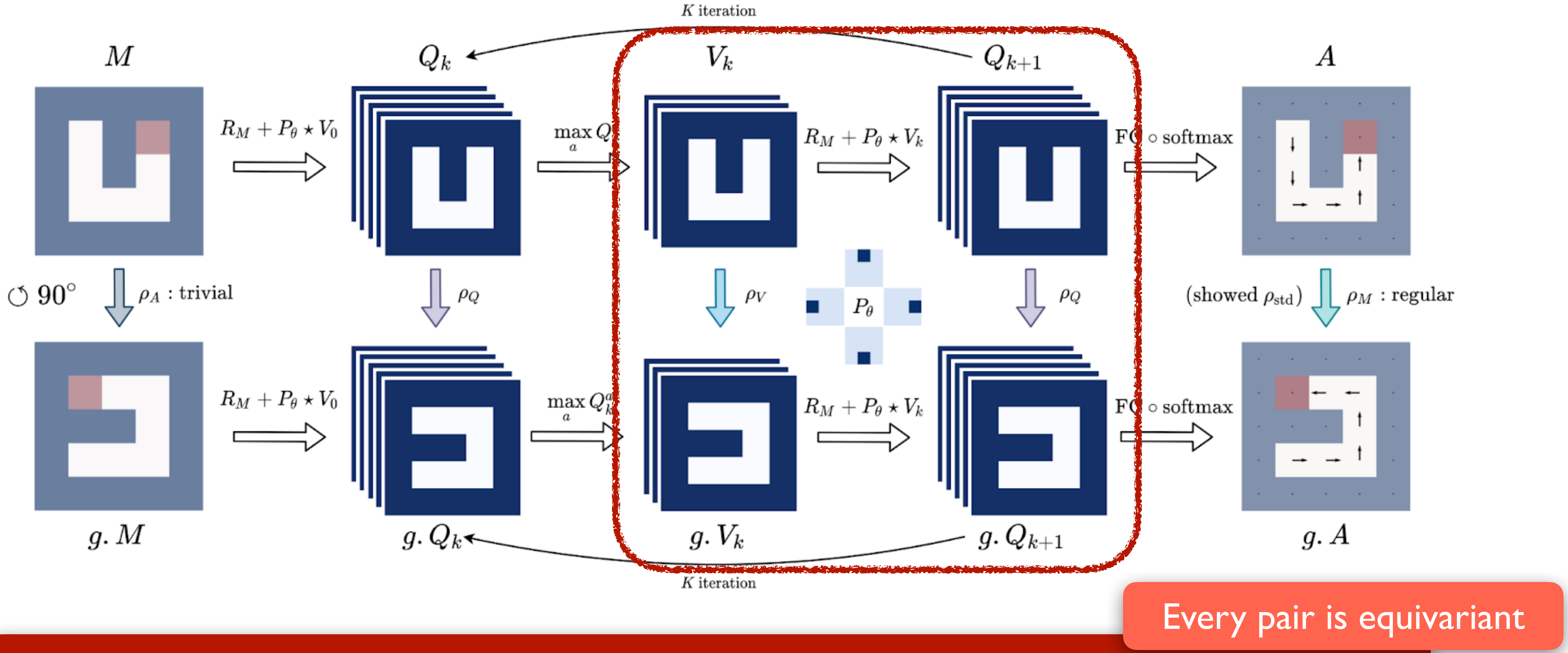
Entire planning is equivariant — Global Equivariance

$$\circlearrowleft 90^\circ \circ \mathrm{VI}(M) \equiv \circlearrowleft 90^\circ \circ \mathcal{T}^\infty[V_0] = \mathcal{T}^\infty[\circlearrowleft 90^\circ \circ V_0] \equiv \mathrm{VI}(\circlearrowleft 90^\circ \circ M)$$

- Use steerable convolution, equivariant to rotation and reflection:

$$\bar{Q}^{(k)} = \bar{R}^a + \mathrm{Conv2D}(\bar{V}^{(k-1)}; W^V_{\bar{a}}) \xrightarrow{\text{Replace}} \bar{Q}^{(k)}_{\bar{a}} = \bar{R}_{\bar{a}} + \mathrm{SteerableConv}(\bar{V}; W^V)$$

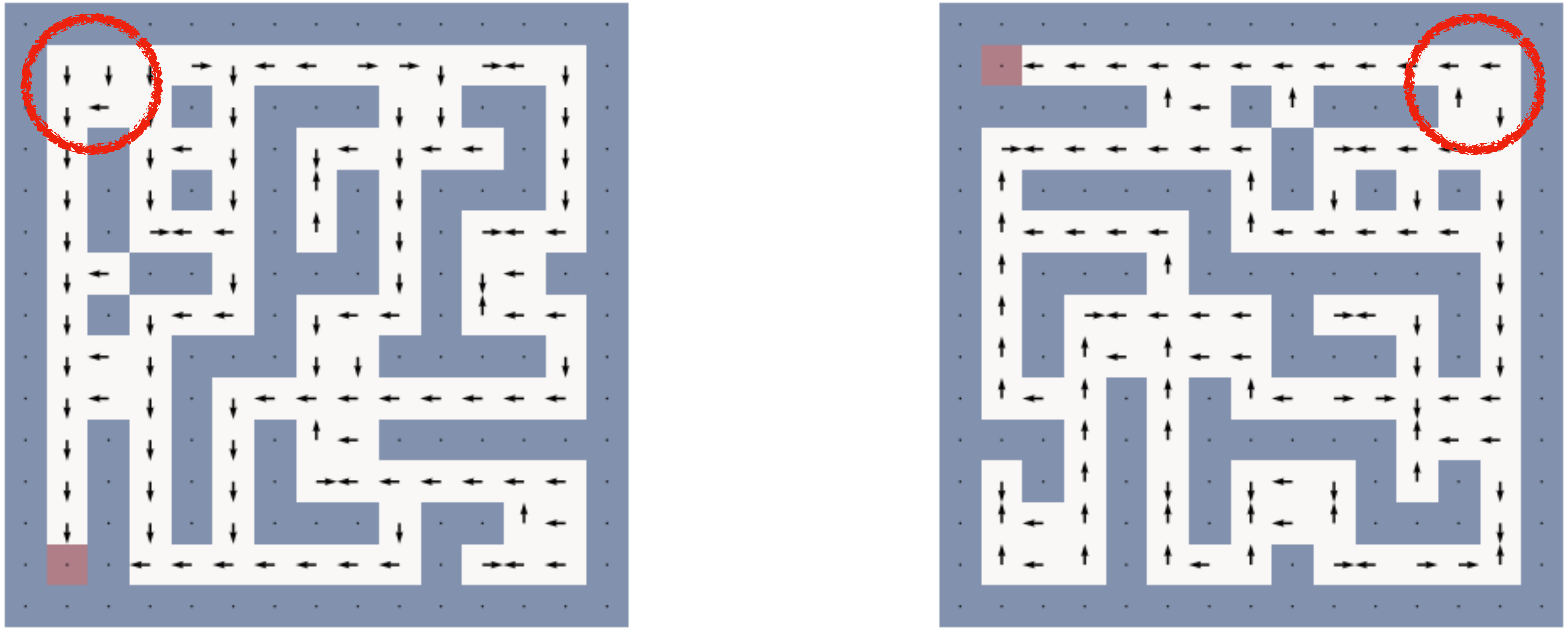
Main Pipeline: Symmetric Value Iteration Network



We use steerable convolutions to integrate symmetry in VINs.

Visualization: VIN

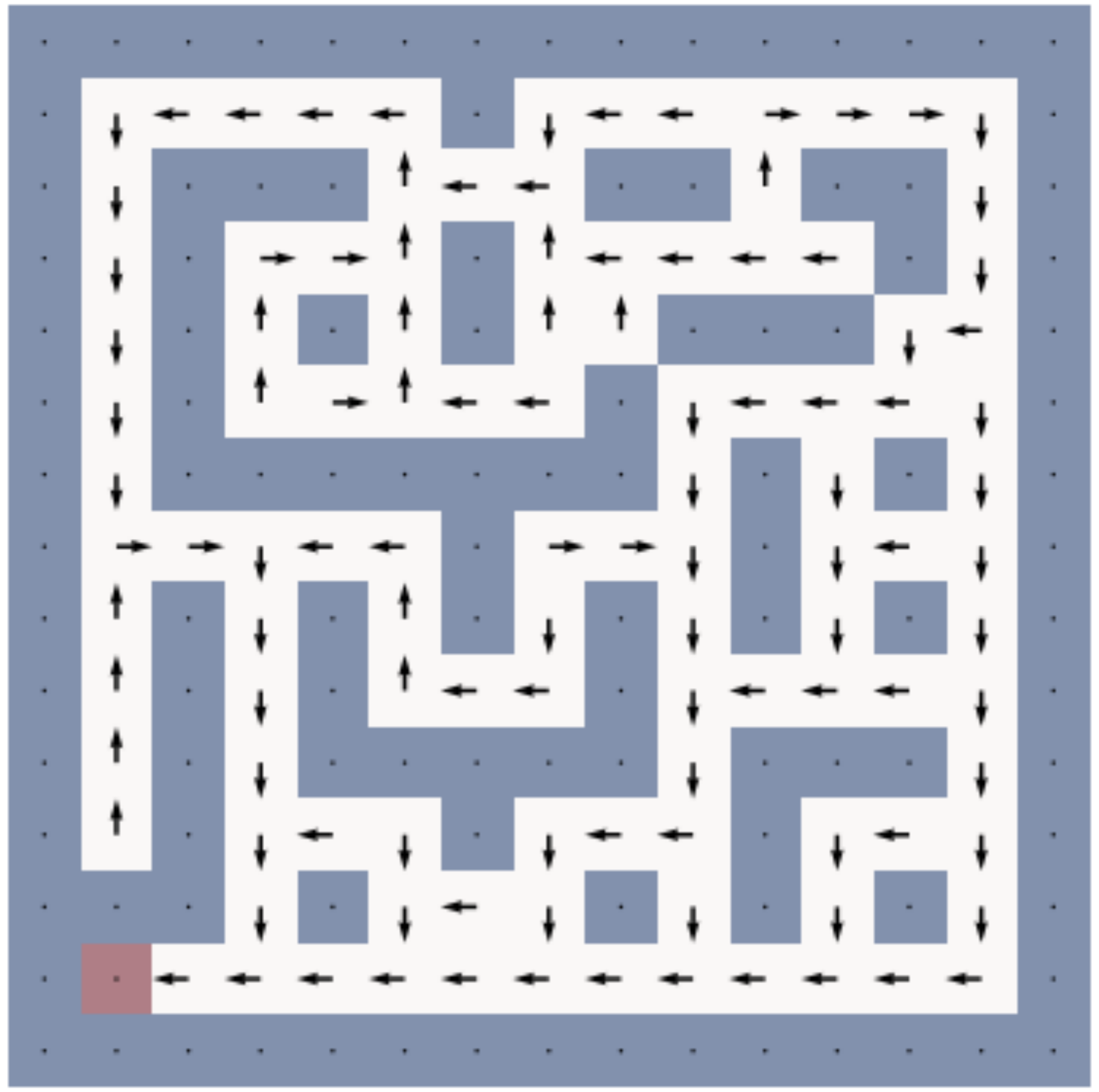
Feed in M and $\circlearrowleft 90^\circ \circ M$

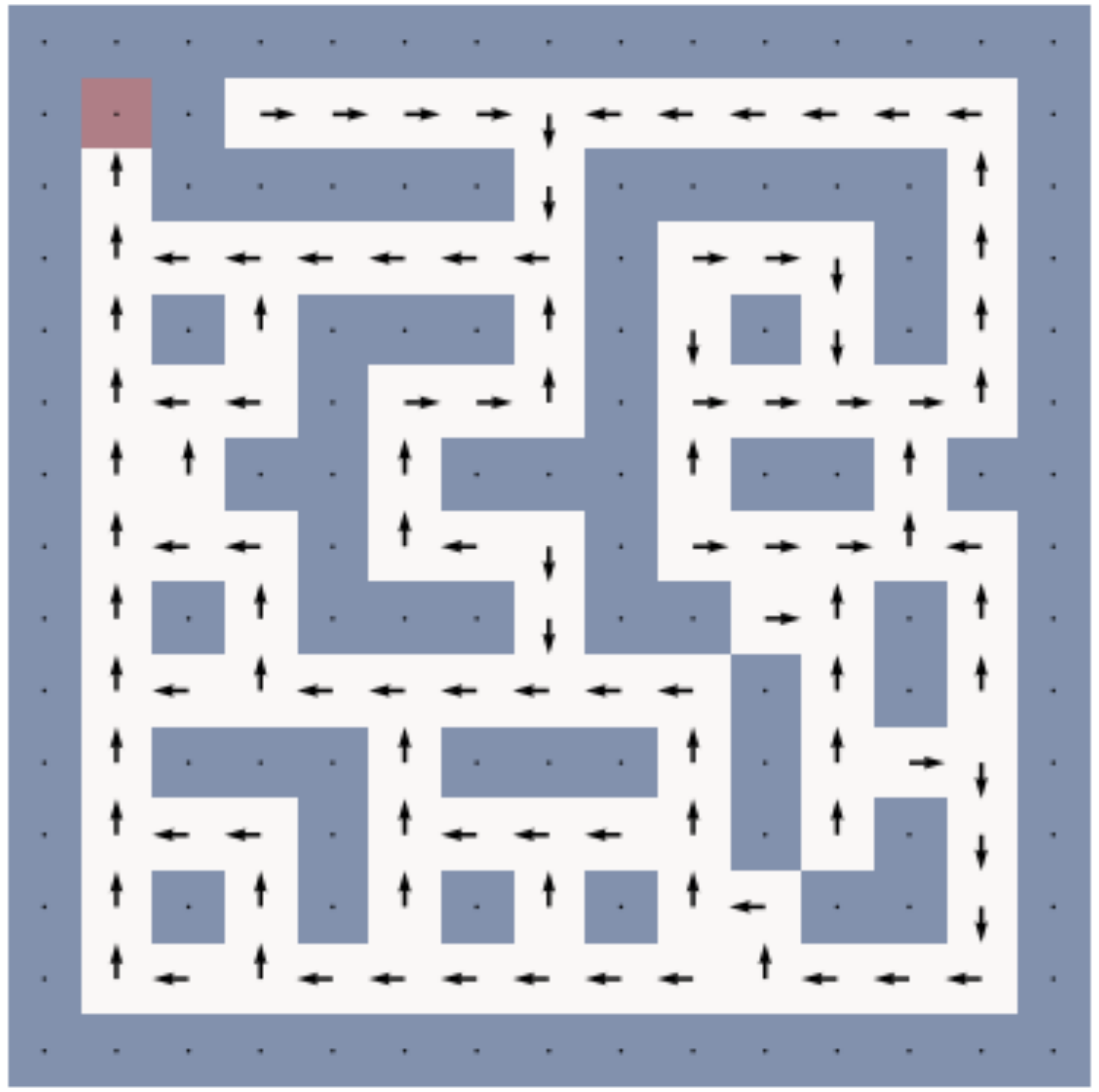


VIN output doesn't satisfy equivariance

Visualization: SymVIN

Feed in M and $\circlearrowleft 90° \circ M$

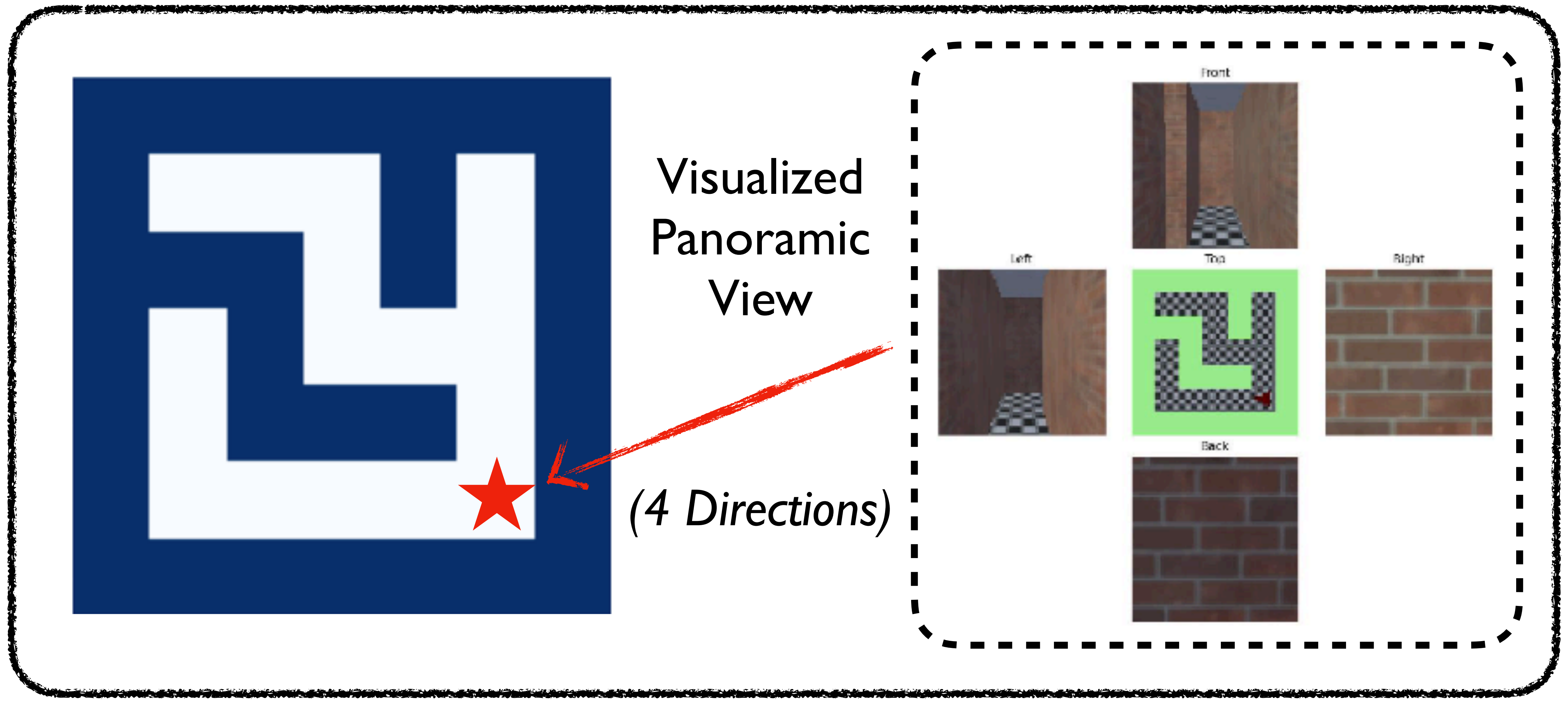


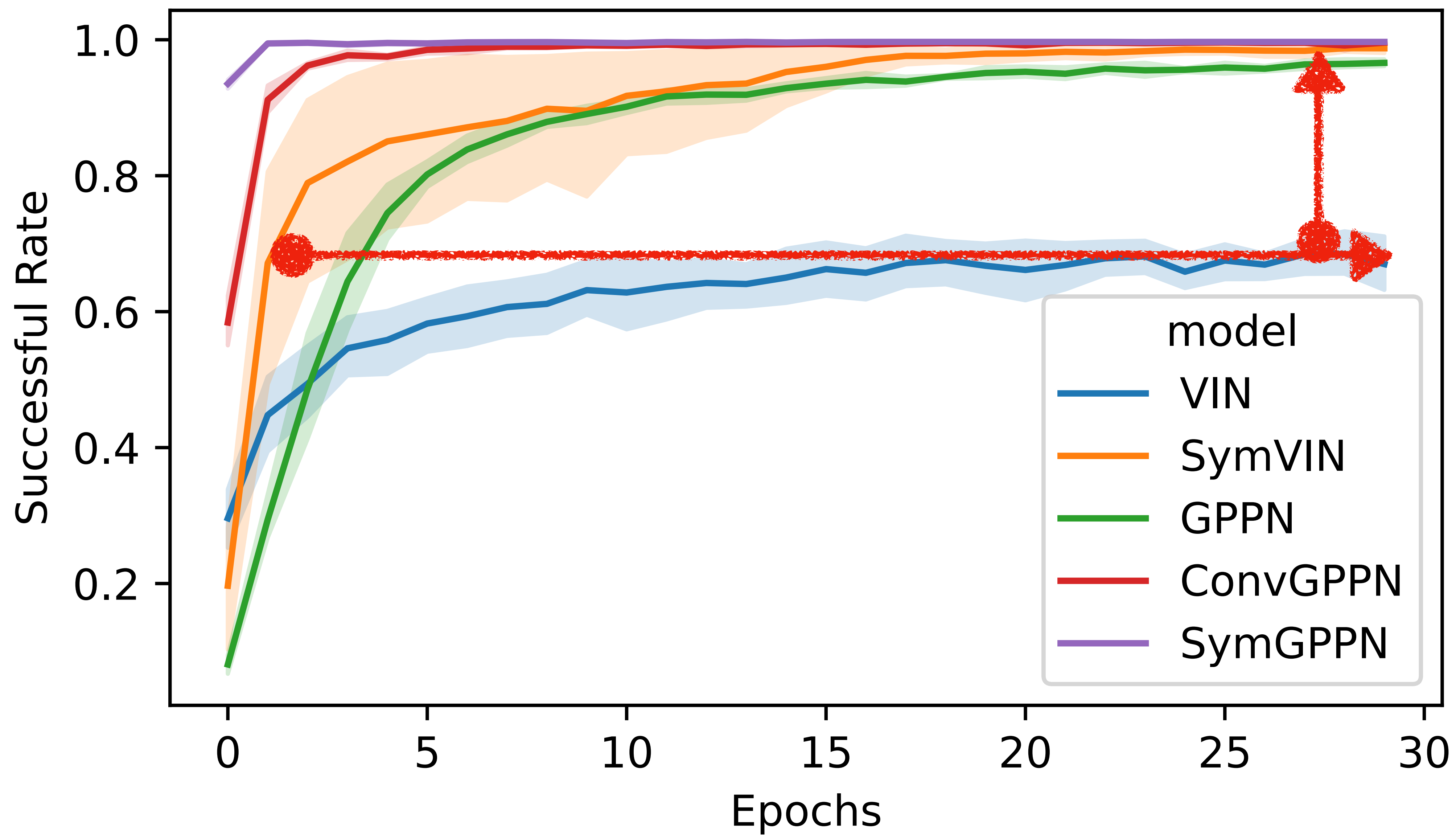


SymVIN guarantees output is equivariant

Experiment: Maze Navigation

2D and Visual Maze Navigation





More efficient training; Higher asymptotic performance

Results: Evaluation on test maps

Method (10K Data)	Navigation 15 × 15	28 × 28	50 × 50	Visual	Manipulation 18 × 18	36 × 36	Workspace
VIN	66.97	67.57	57.92	50.83	77.82	84.32	80.44
SymVIN	98.99	98.14	86.20	95.50	99.98	99.36	**91.10**
GPPN	96.36	95.77	91.84	93.13	2.62	1.68	3.67
ConvGPPN	99.75	99.09	97.21	98.55	99.98	99.95	89.88
SymGPPN	**99.98**	**99.86**	**99.49**	**99.78**	**100.00**	**99.99**	90.50

- Better generalization on novel maps
- Test novel maps are not necessarily rotated version of training maps

Theoretical results

Theorem 1 (informal): **Value iteration** for **path planning*** is equivariant to translation, rotation, and reflection

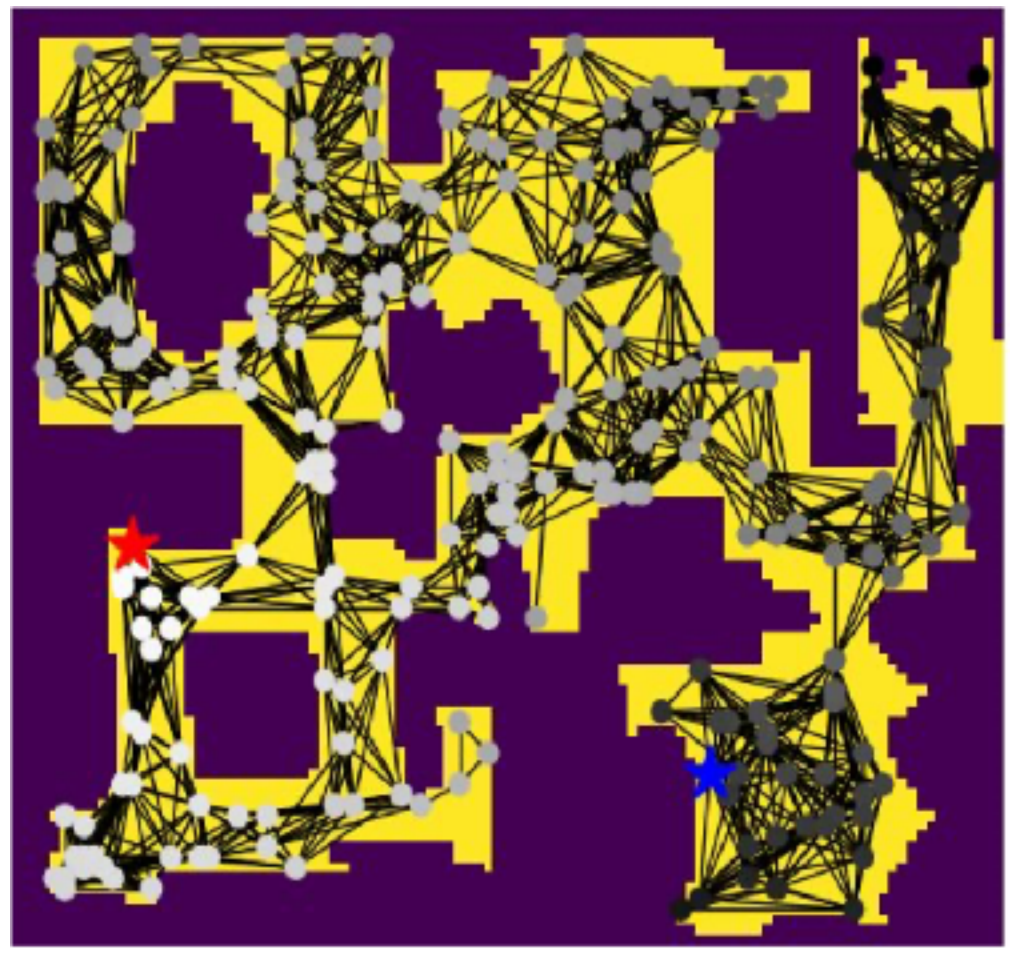
Theorem 2 (informal): **Value iteration** for **path planning*** is a form of **steerable convolution network****

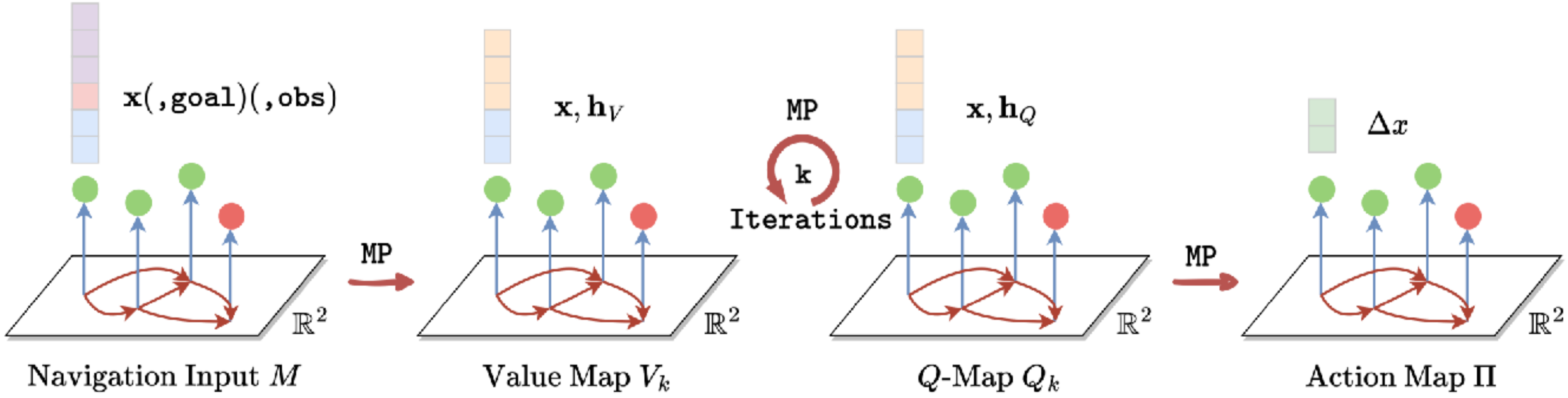
*: Path planning on 2D grid, an example of homogeneous spaces
**: Steerable CNN over grids, equivariant under induced representations

Cohen et al. (2017): Steerable CNNs, ICLR 2017

Follow-up: Path Planning on Graphs

We extend path planning with 2D convolution to graph convolution and message passing layers on graphs.





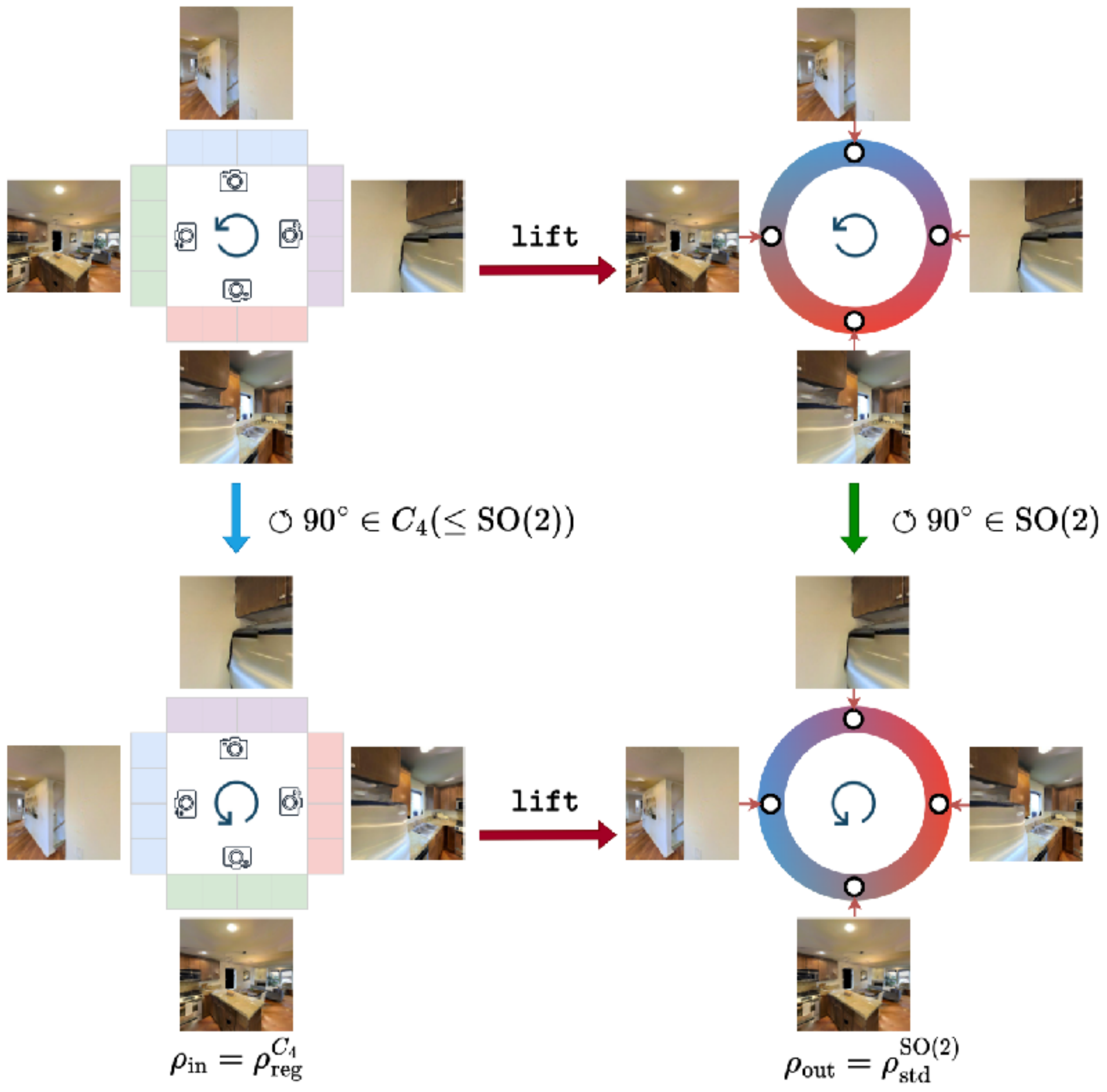
Zhao*, Li*, Padir, Jiang†, Wong†. *"E(2)-Equivariant Graph Planning for Navigation"*.
RA-L 2023 & IROS 2024 (Oral).

Challenge: Camera/View Layout

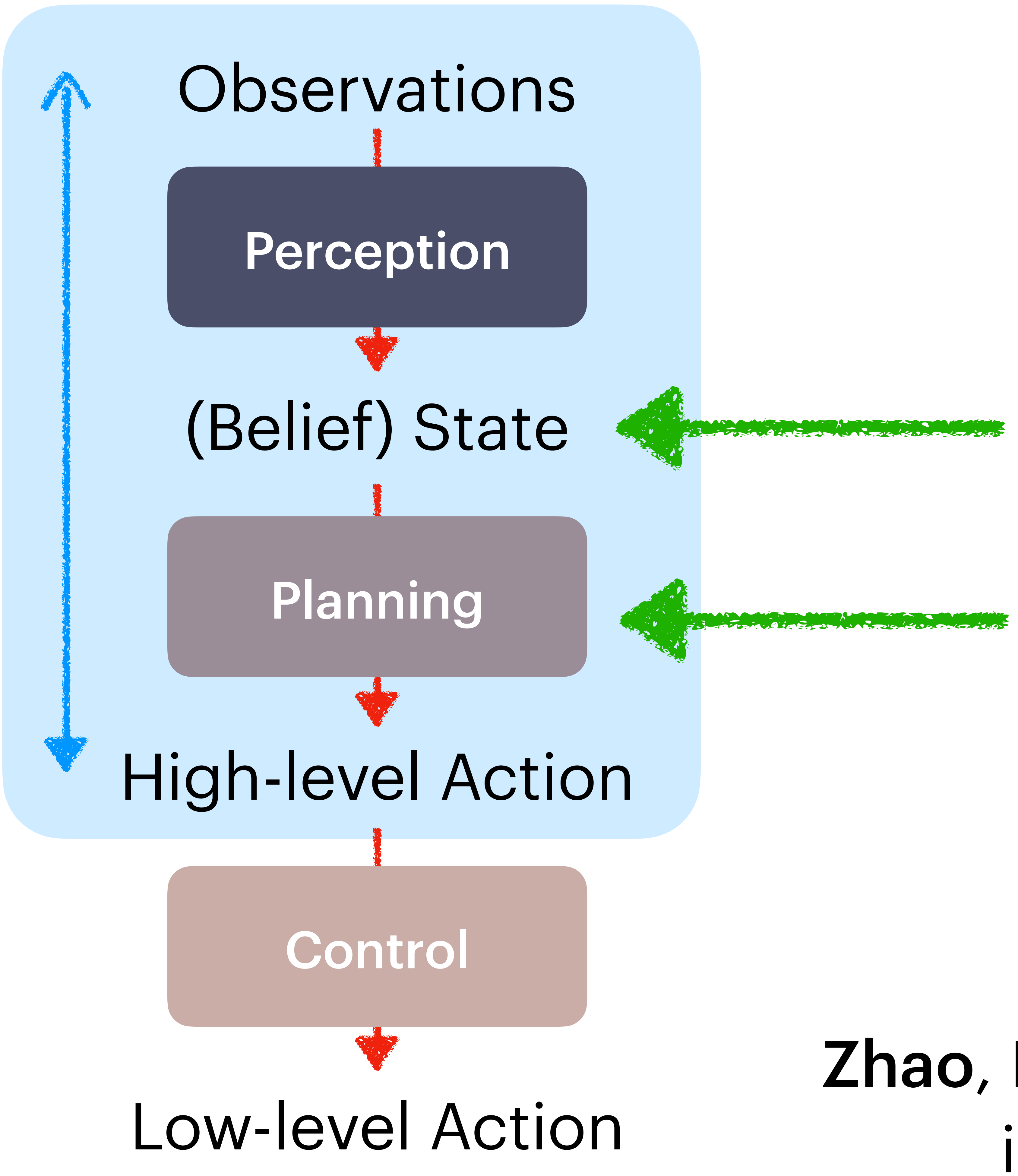
Robots may only have K views

- Naive equivariance only allow C_K (or $\frac{360°}{K}$) rotation symmetry
- We lift it to $\mathrm{SO}(2)$ to allow *continuous* symmetry in downstream planning network

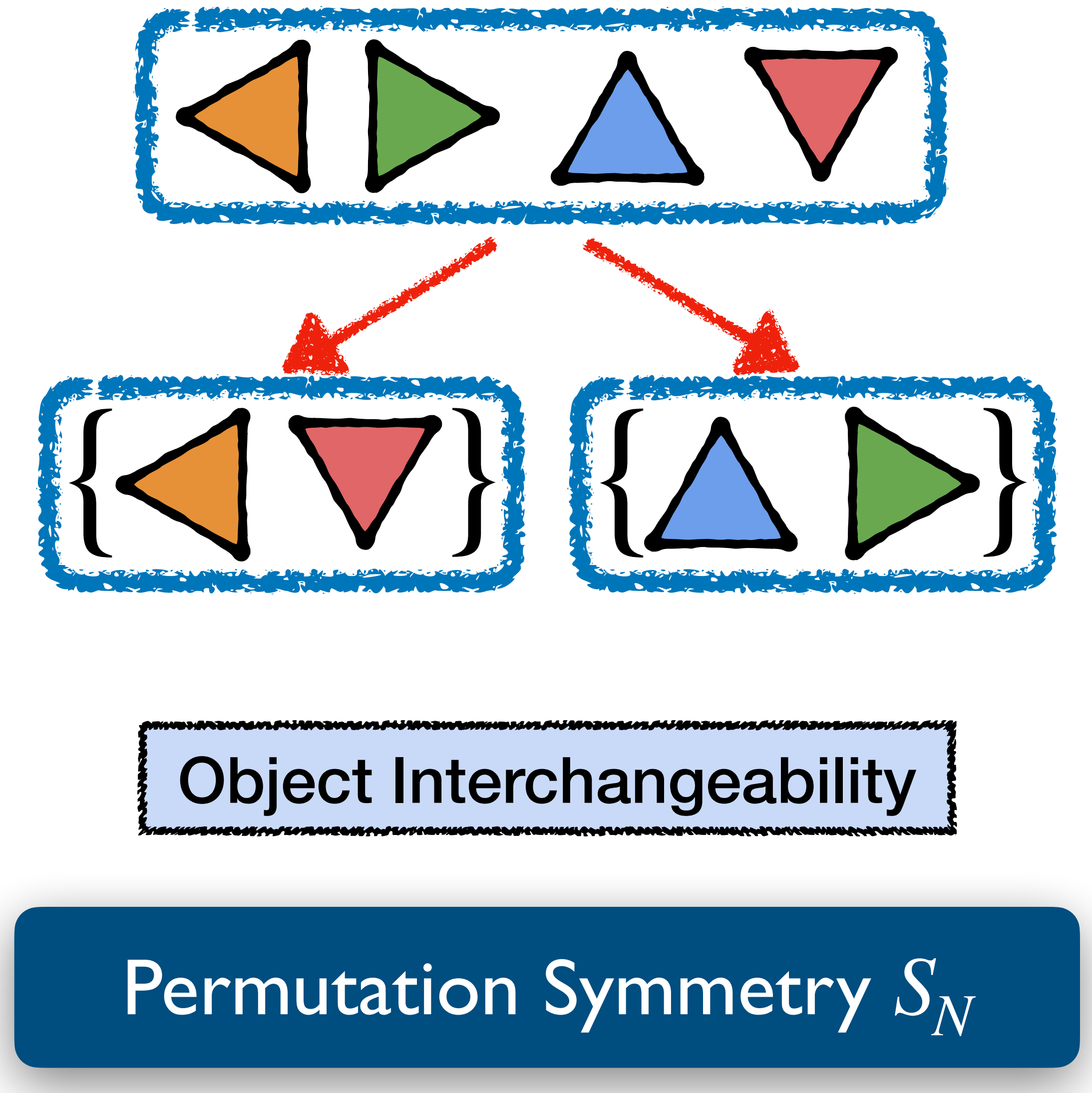
Commutative diagram of the ***lift*** layer:



Object Compositionality in World Modeling



When object structure presents, the state space may be factored into e.g., slots.

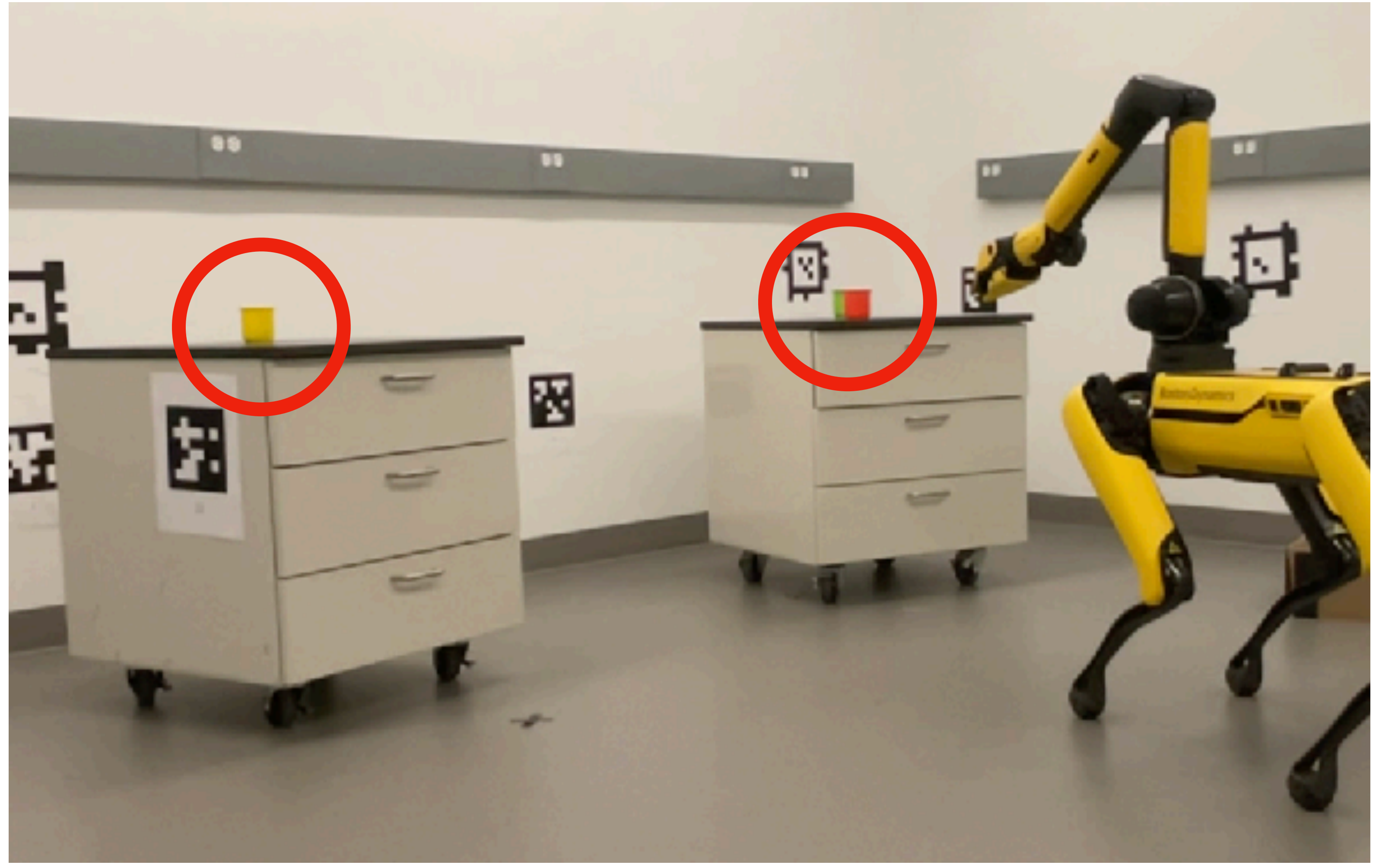


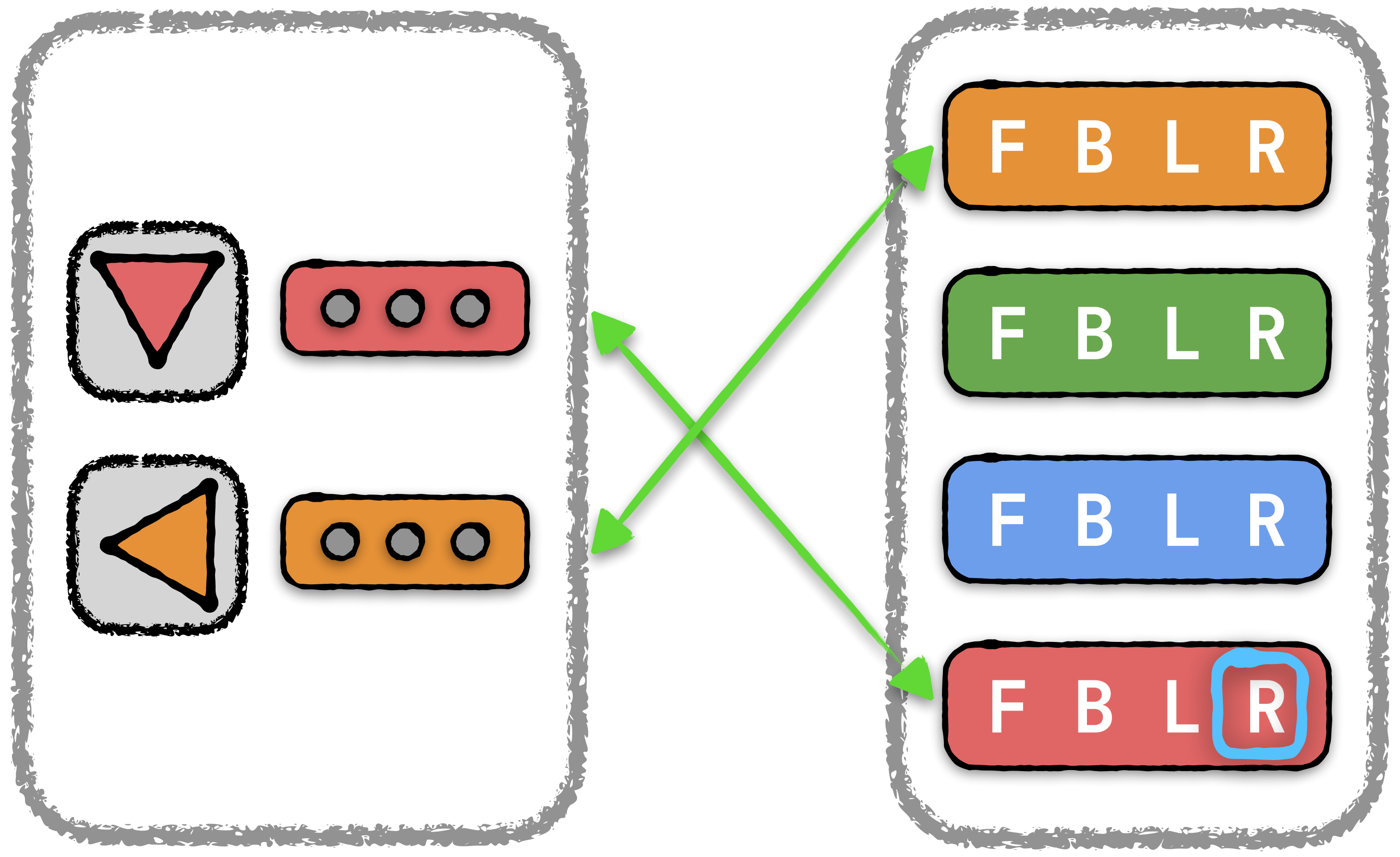
This implicitly produces a reduced MDP. How to represent and plan in this MDP while keep differentiable?

Zhao, Kong, Walters, Wong. "Toward Compositional Generalization in Object-Oriented World Modeling". ICML 2022 (**Oral**).

Key Ideas

Object Compositionality in World Modeling and Planning

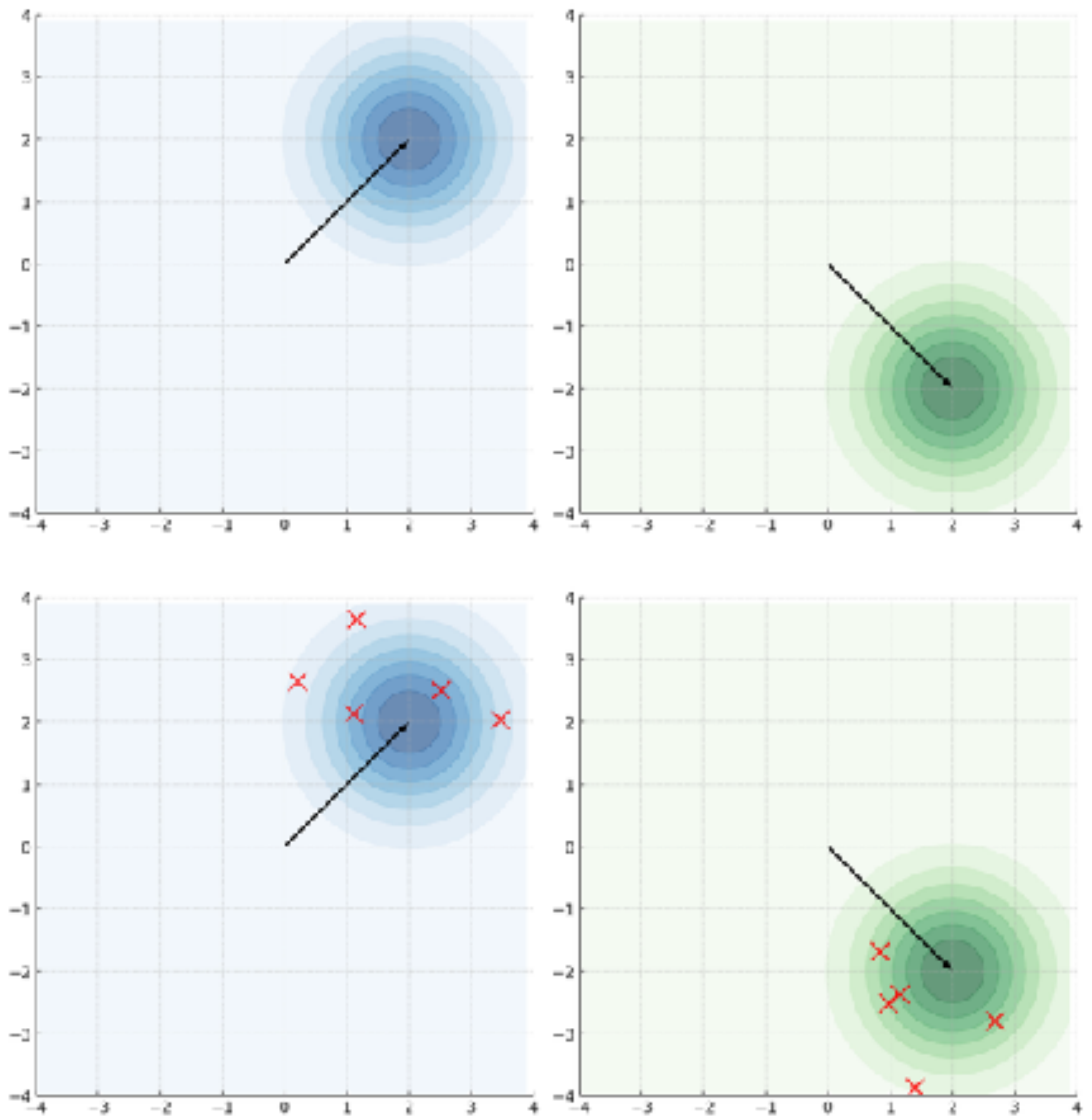




- Objects don't have order / universally unique identifier
- When learning a model and planning, actions and slots need to ***bind*** correctly to the desired objects
- Correct binding provably induces a smaller "slot MDP" for more efficient planning

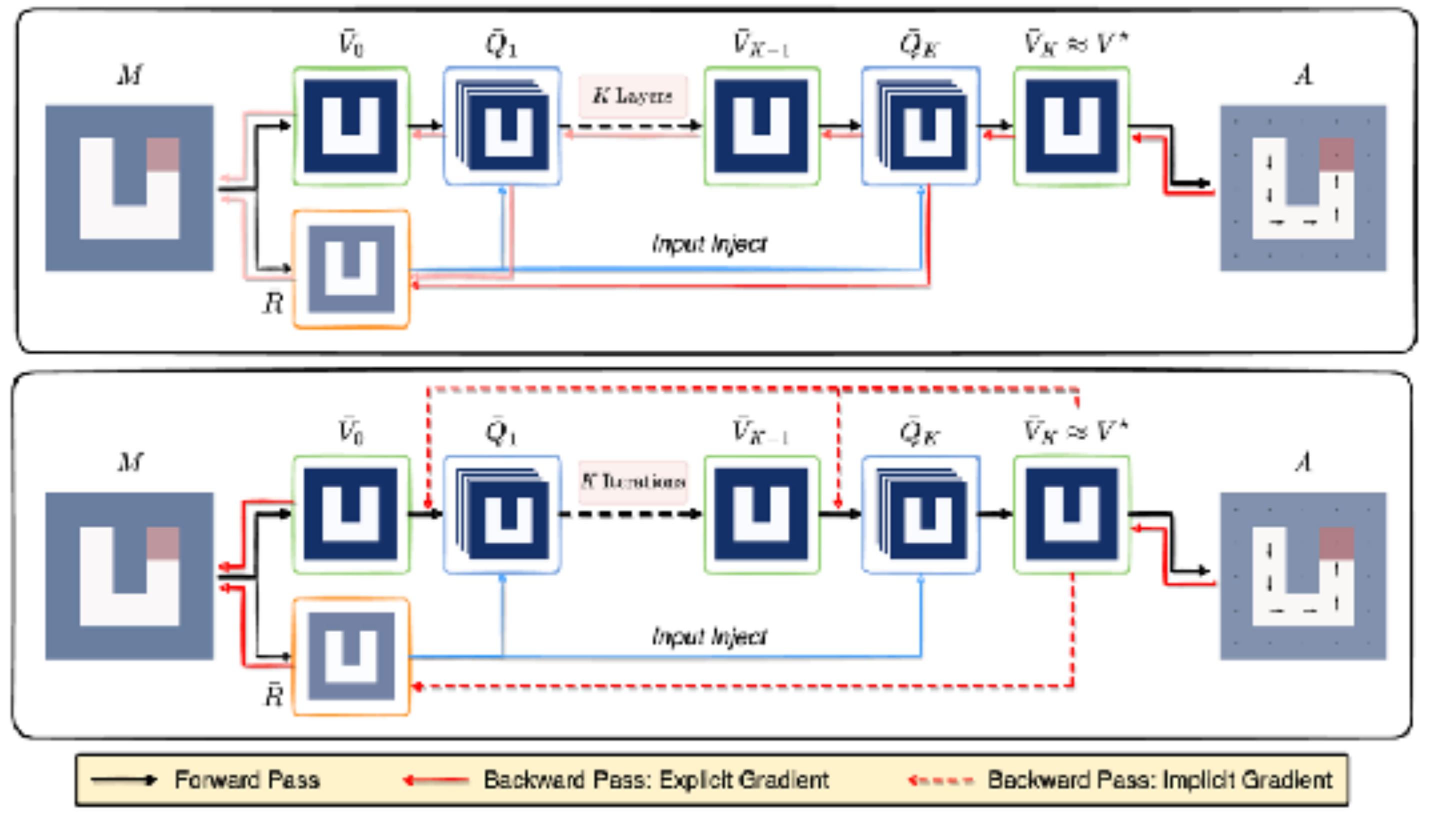
Other Work

On Structured Learning, Lossless Representation and Extension



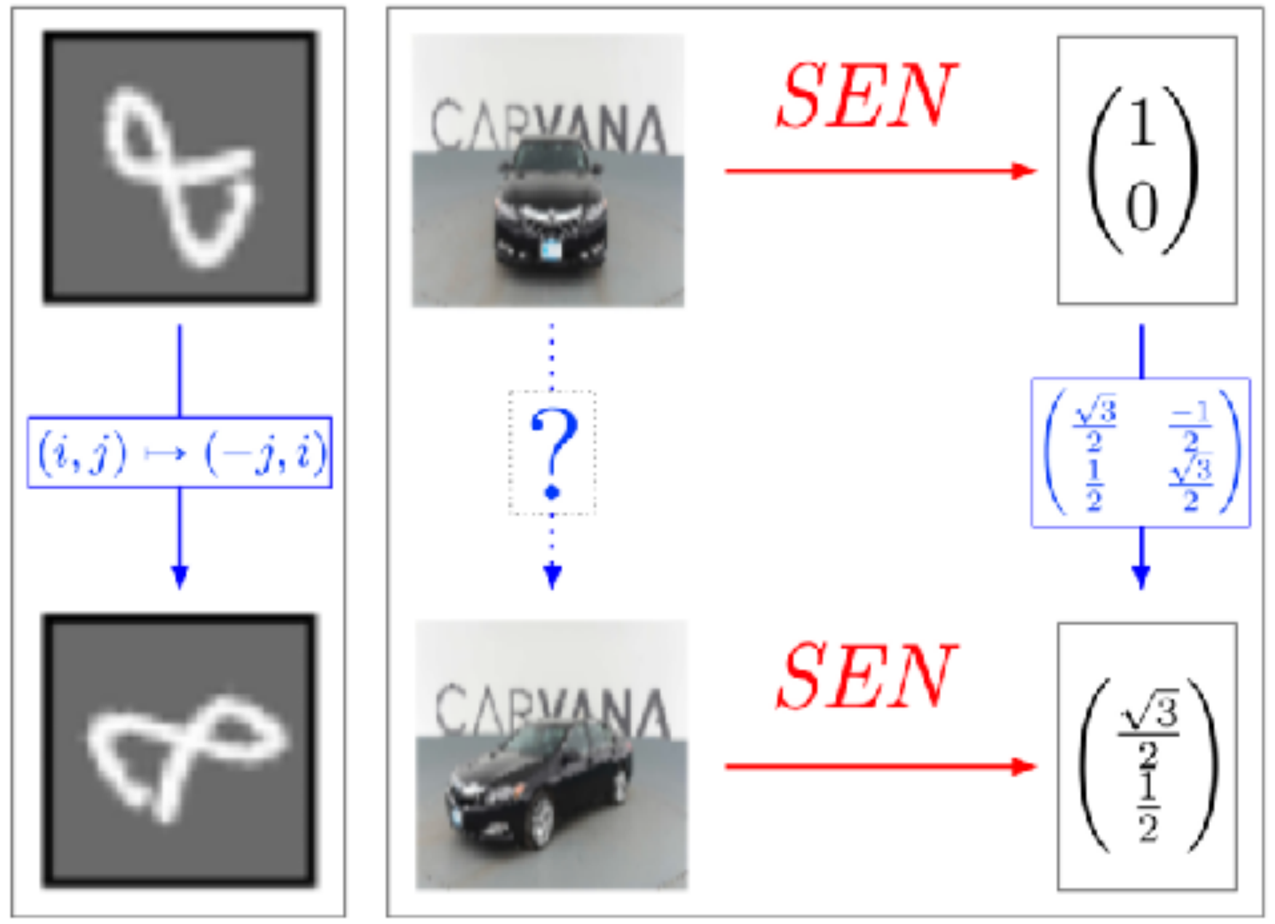
Equivariant Sampling

Zhao, Howell, Zhu, Park, Zhang, Walters†, Wong†.
WAFR 2024.



Implicit Differentiation for Planning

Zhao, Xu, Wong.
ICLR 2023.



Symmetric Representation

Park*, Biza*, **Zhao**, van de Meent, Walters.
ICML 2022.

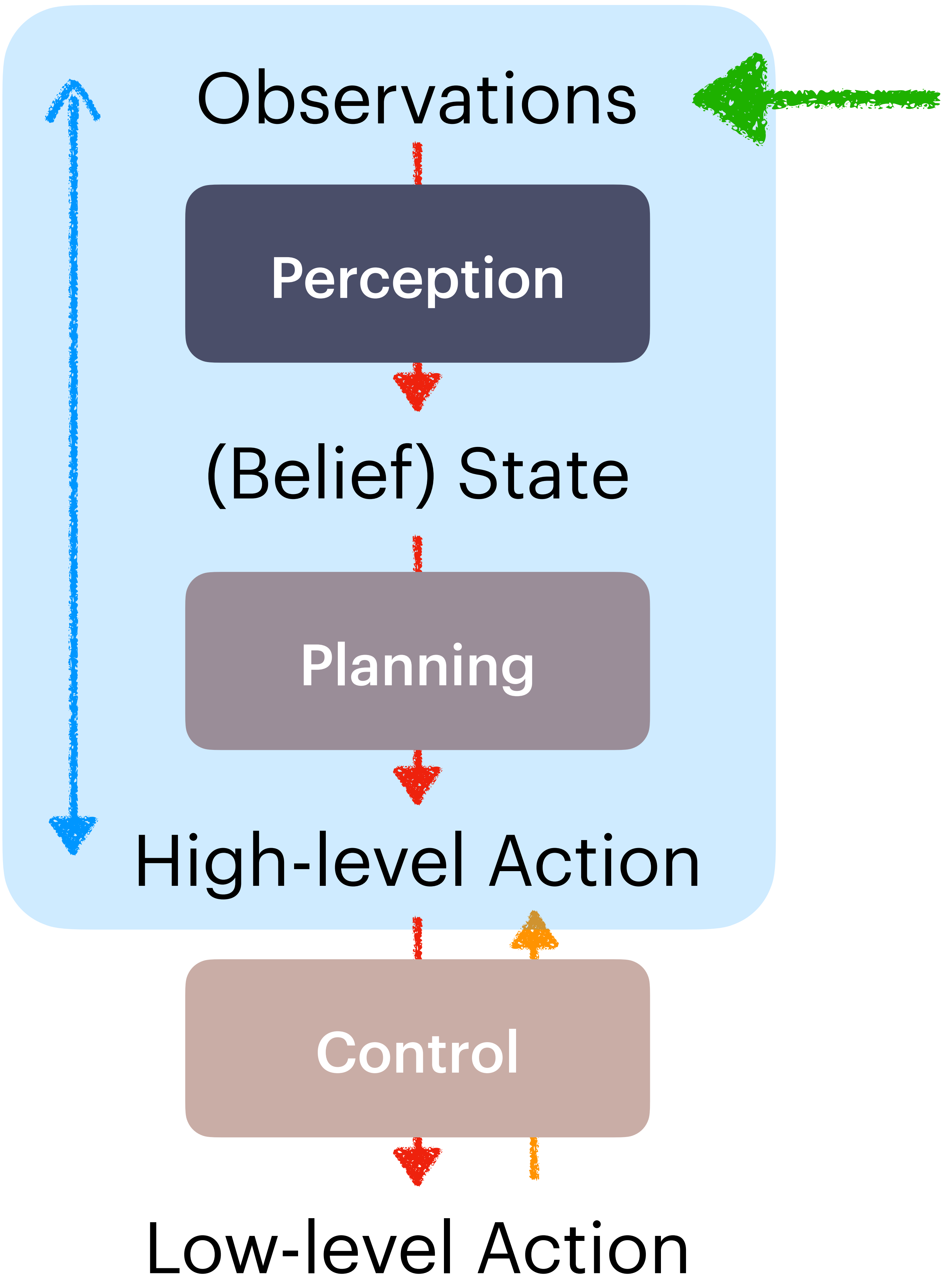
Summary

Lossless Abstraction of World Representation and Planning

- ✅ By retaining all critical information from the environment, lossless abstractions allow for exact and high-fidelity planning.
- ⚠️ Lossless representations environment can be computationally intensive due to high-dimensional state and action spaces.
    - ⚠️ Applying these methods in large-scale, dynamic environments is challenging.

Part 2: *Lossy* Abstraction of World Representation and Planning

Real-world Lossless representation + Planning is Too Hard

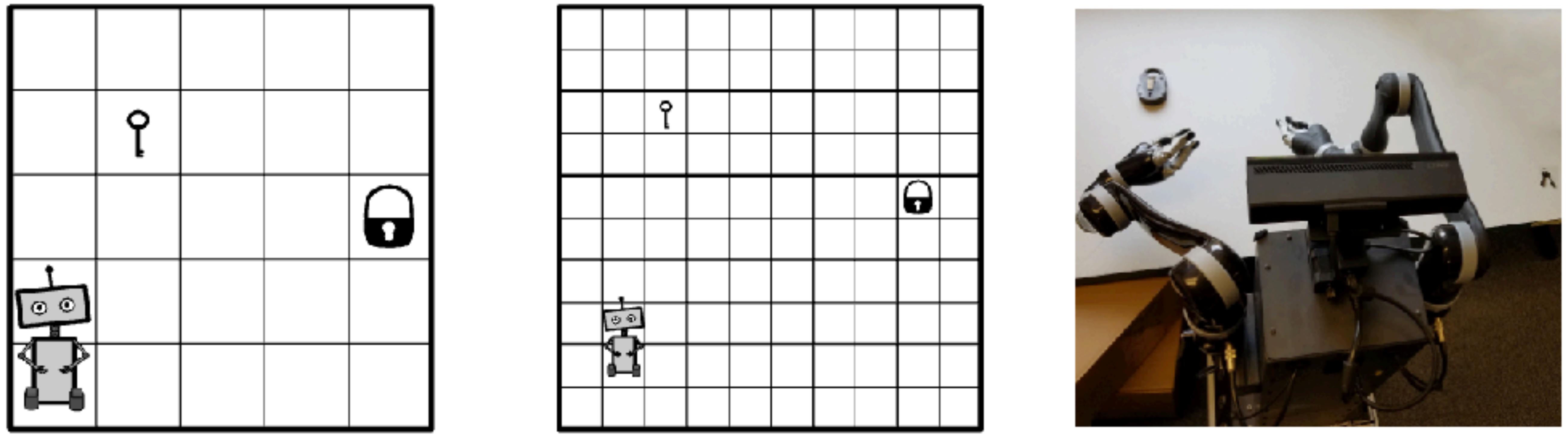


Spot has 6 cameras (5 body + 1 in-hand)

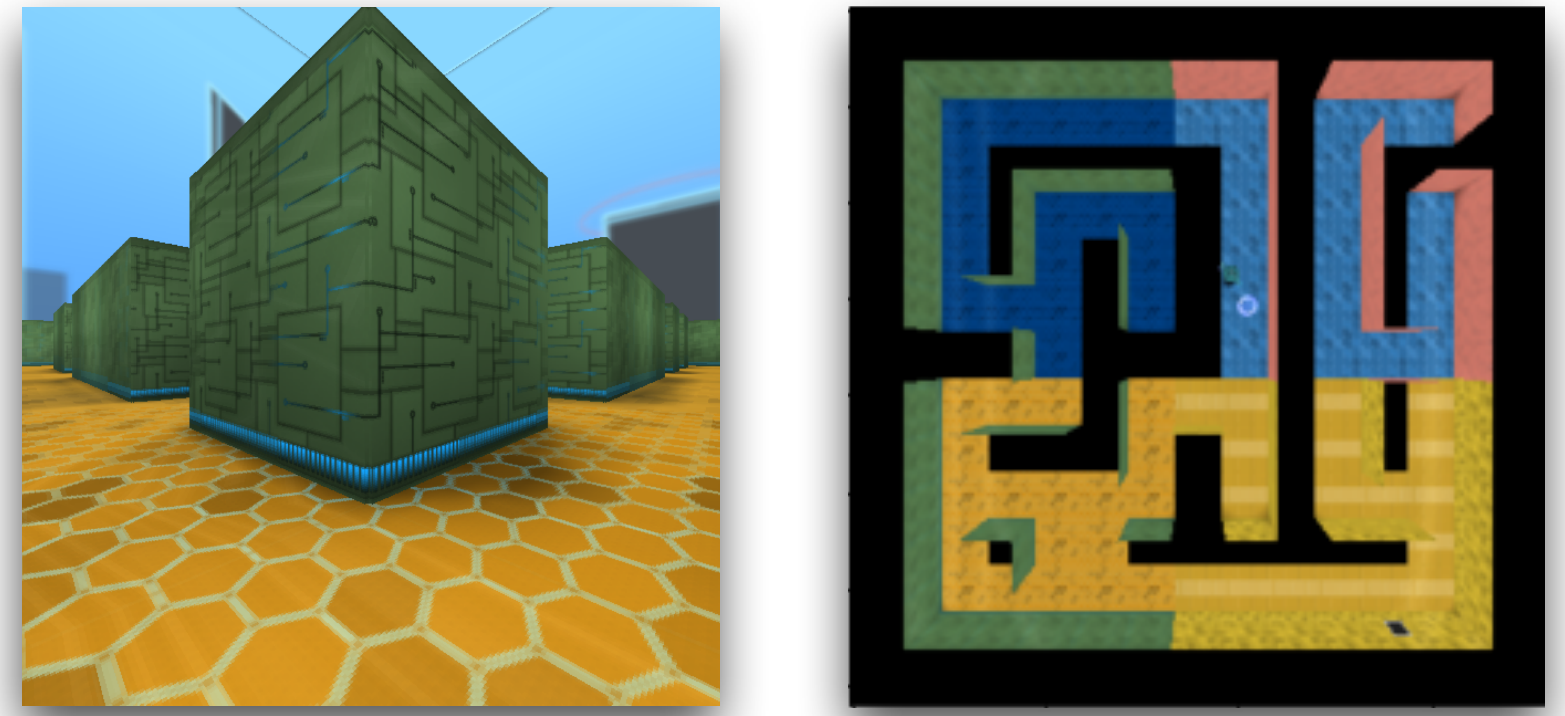
We have to abandon some details!

"Lossy" Abstraction

Examples



Symbolic/Language-based Representation
Planning: Options/Operators/Skills



Abstract Maps
Planning: 2D Paths

Konidaris et al. From Skills to Symbols: Learning Symbolic Representations for Abstract High-Level Planning. IJRR 2018.
Silver et al. Learning Neuro-Symbolic Skills for Bilevel Planning. CoRL 2022.
Xu et al. Robot Navigation in Unseen Environments using Coarse Maps. ICRA 2024.

“Lossy” Abstraction

Examples: Language

Autonomous Robots (2019) 43:449–468
https://doi.org/10.1007/s10514-018-9792-8

Grounding natural language instructions to semantic goal representations for abstraction and generalization

Dilip Arumugam[1] · Siddharth Karamcheti[1] · Nakul Gopalan[1] · Edward C. Williams[1] · Mina Rhee[1] · Lawson L. S. Wong[1] · Stefanie Tellex[1]

Dilip Arumugam

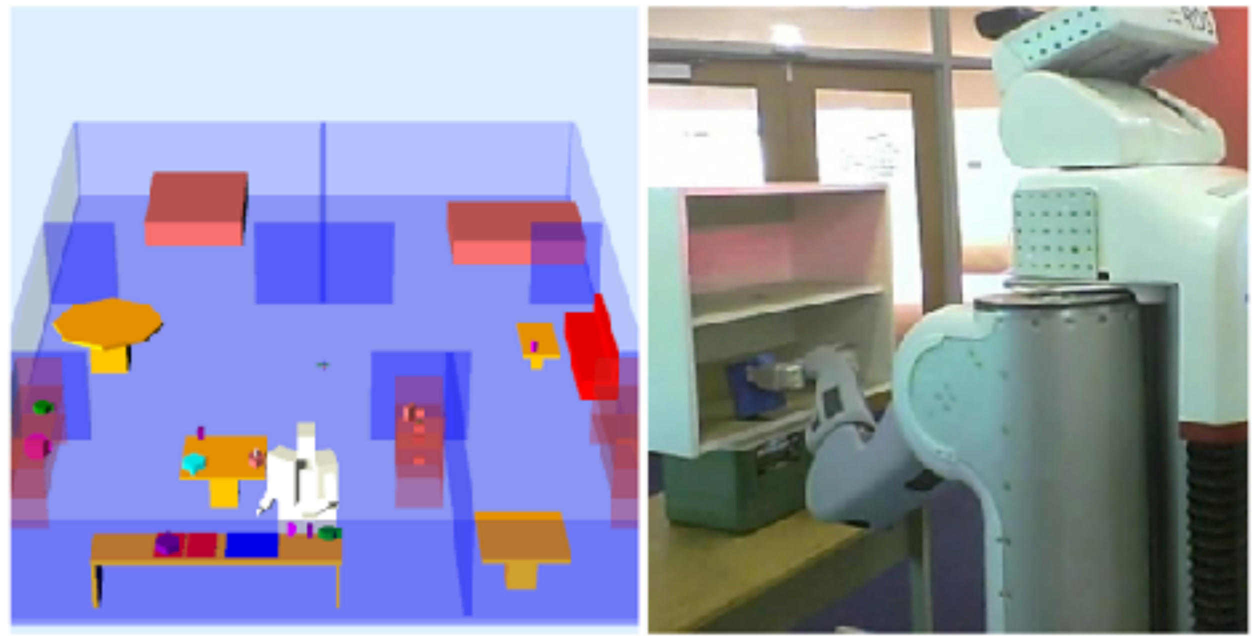
Siddharth Karamcheti

Nakul Gopalan

Stefanie Tellex

[RSS 2017, RSS 2018, AURO 2019]

“Lossy” Abstraction



- It typically involves ***hierarchical*** structure.
- The high-level representation is typically engineered to be highly abstracted from details.
  - Example: symbolic representation. For ***efficient*** planning, they typically abandon details, e.g., geometric features.
- Planners need to be aware of the abstraction and ***ground*** abstract actions.

Leslie Pack Kaelbling and Tomás Lozano-Pérez, Hierarchical Planning in the Now. ICRA 2011.
Leslie Pack Kaelbling and Tomás Lozano-Pérez, Integrated Task and Motion Planning in Belief Space, *International Journal of Robotics Research*, 2013

Symbolic Planning with Parameterized Skills

Bilevel Planning

Idea:

- Build a high-level symbolic model
- Hand-design skills+operators
- Use AI planner to solve high-level planning problem
- Then ground symbolic actions to physical world

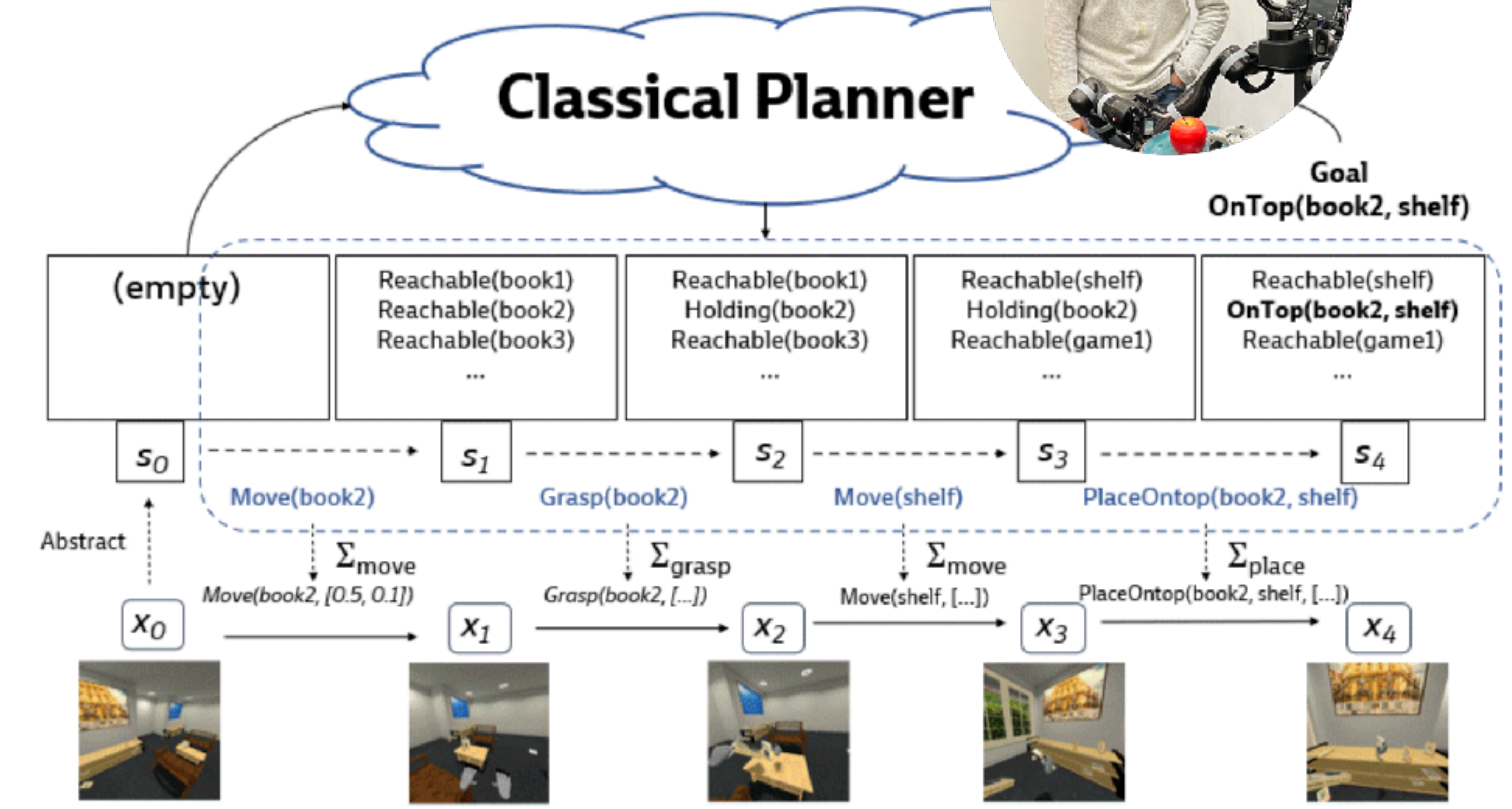
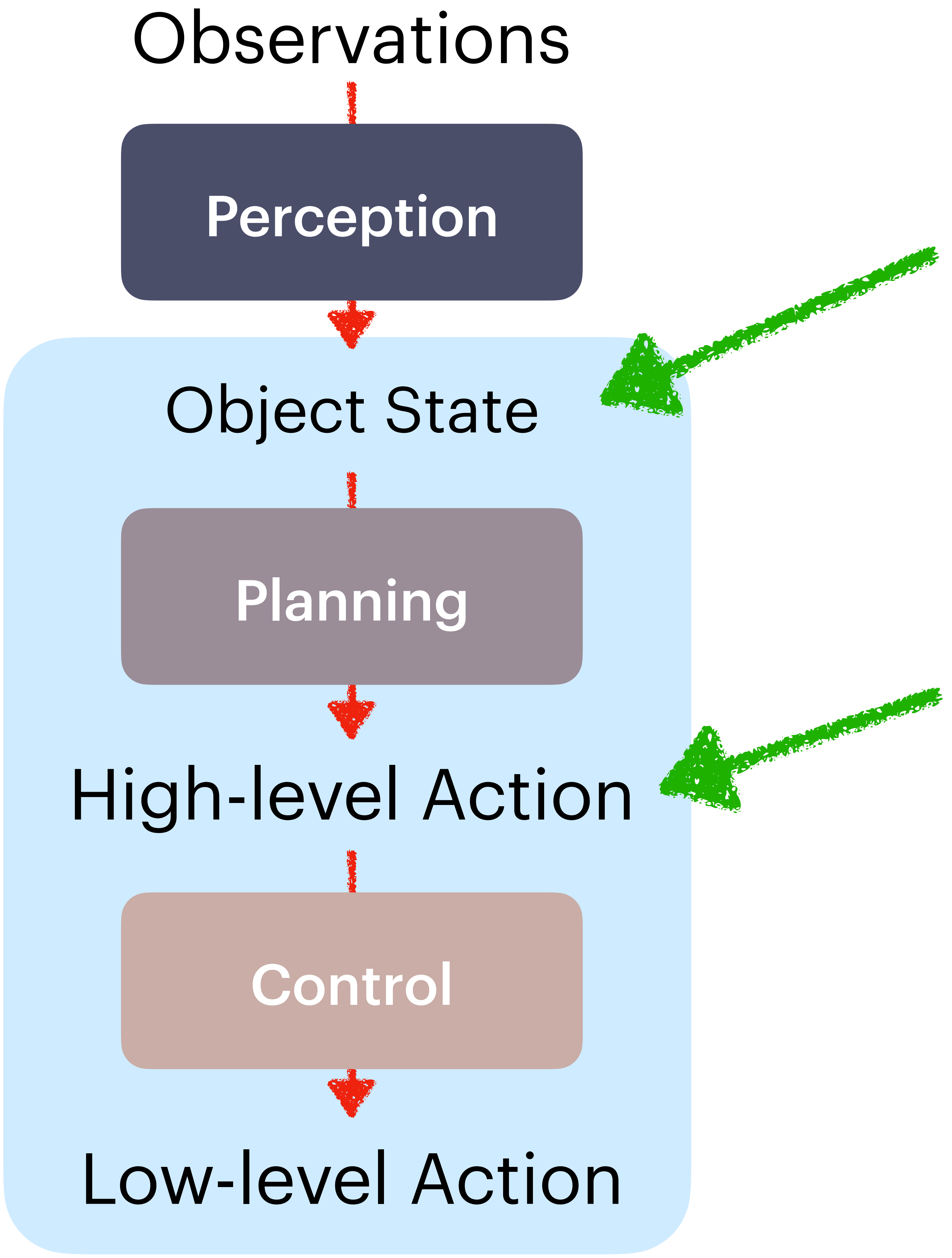


Figure 8: Animated visualization of constructing an abstract plan, and then 'refining' this plan using samplers (denoted by ∑) to derive the continuous parameters for skill associated with an operator. These skills now have all their parameters specified, so can be executed in the environment in sequence.

Credit: Kumar, McClinton, Silver et al.

Planning to Practice Skill Parameters



Planning in abstract symbolic space enables long-horizon behaviors

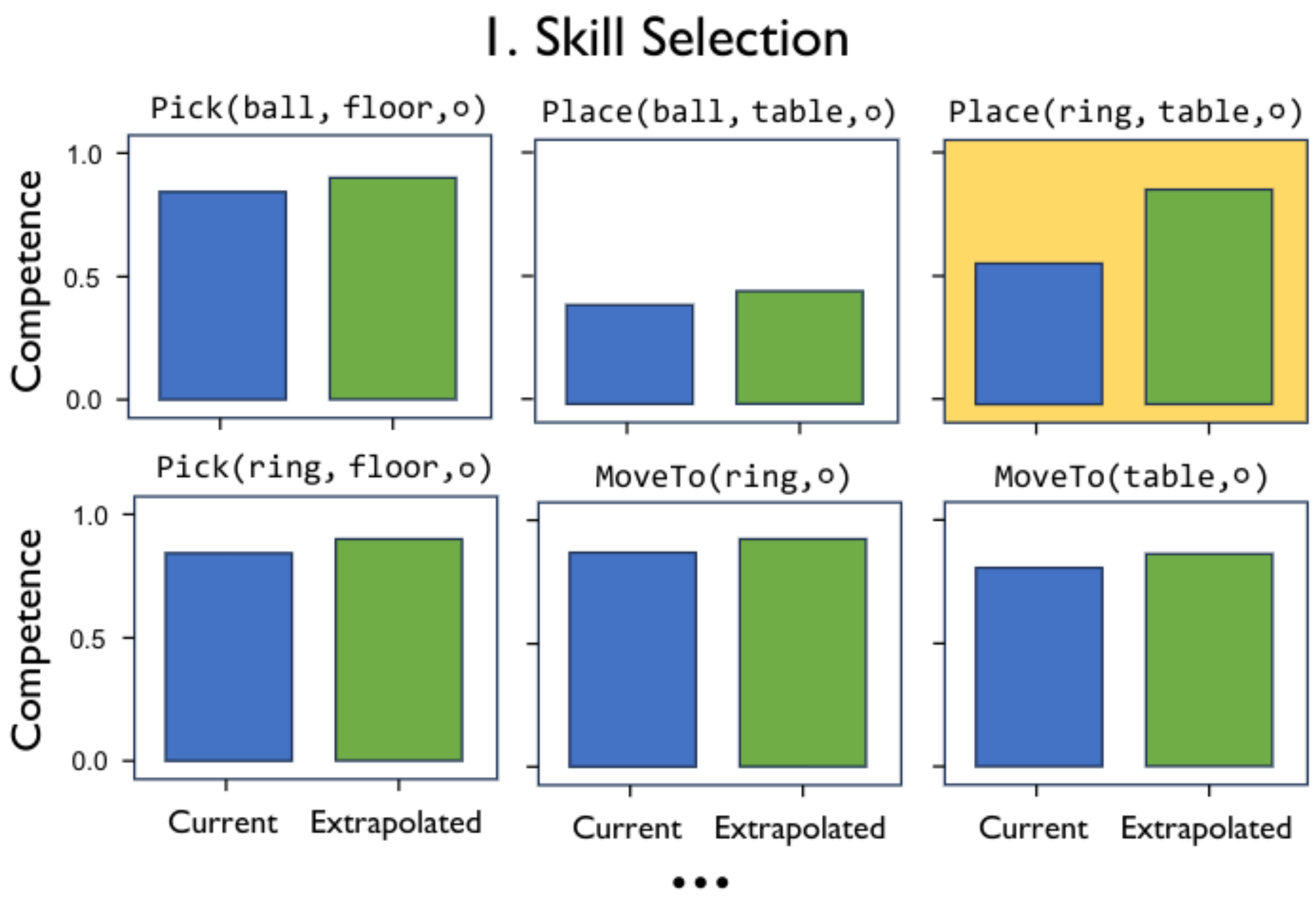
When planning in abstract symbolic space, it couldn't provide physically grounded actions

With symbolic planners, how do we ground them?

Kumar*, Silver*, McClinton, **Zhao**, Proul, Lozano-Pérez, Kaelbling, Barry. *"Practice Makes Perfect: Planning to Learn Skill Parameter Policies"*. RSS 2024.

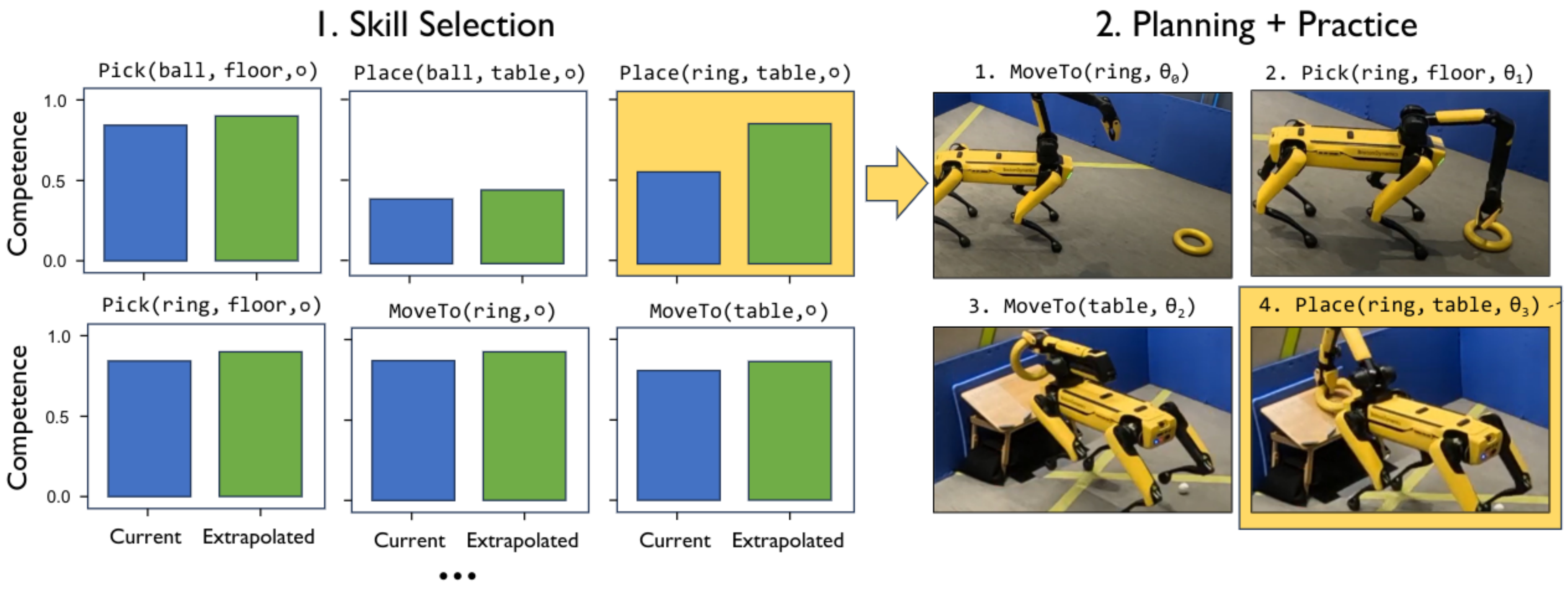
Idea: Planning to Practice Skills

Practice Skills that Expect Most Improvement via Sampling!



Idea: Planning to Practice Skills

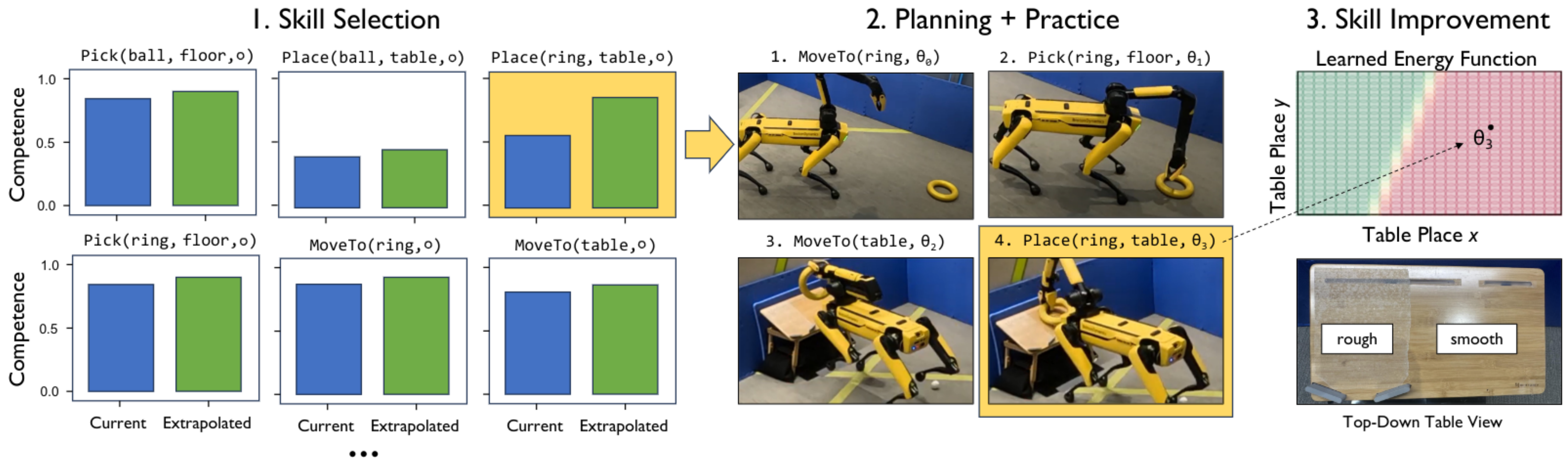
Practice Skills that Expect Most Improvement via Sampling!



Planning enables long-horizon behaviors, allowing practicing skills' parameters

Idea: Planning to Practice Skills

Practice Skills that Expect Most Improvement via Sampling!



Planning enables long-horizon behaviors, allowing practicing skills' parameters

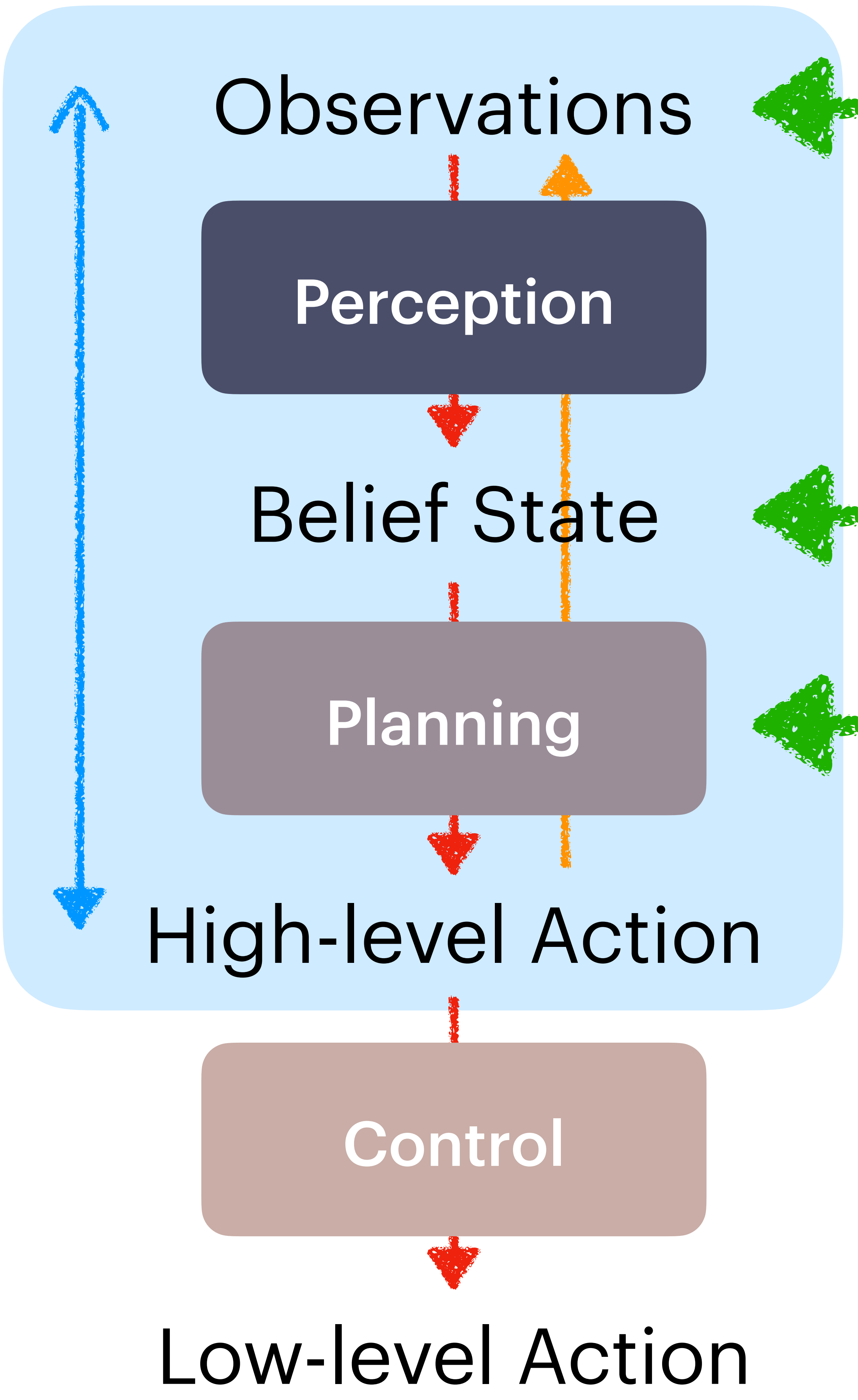

8X
Boston Dynamics

Takeaways

Planning to Practice Parameterized Skills

- ✅ Planning in efficient symbolic representation enables long-horizon robot tasks in real-world
- ✅ ***Real-robot interactions*** are needed for ***grounding*** a symbolic plan into a physically plausible plan when no simulator is available
- ⚠️ The planner assumes **full observability** and **complete knowledge** about the initial state of the world — Strong assumption for mobile robots

Integrating Perception and Belief-space Planning

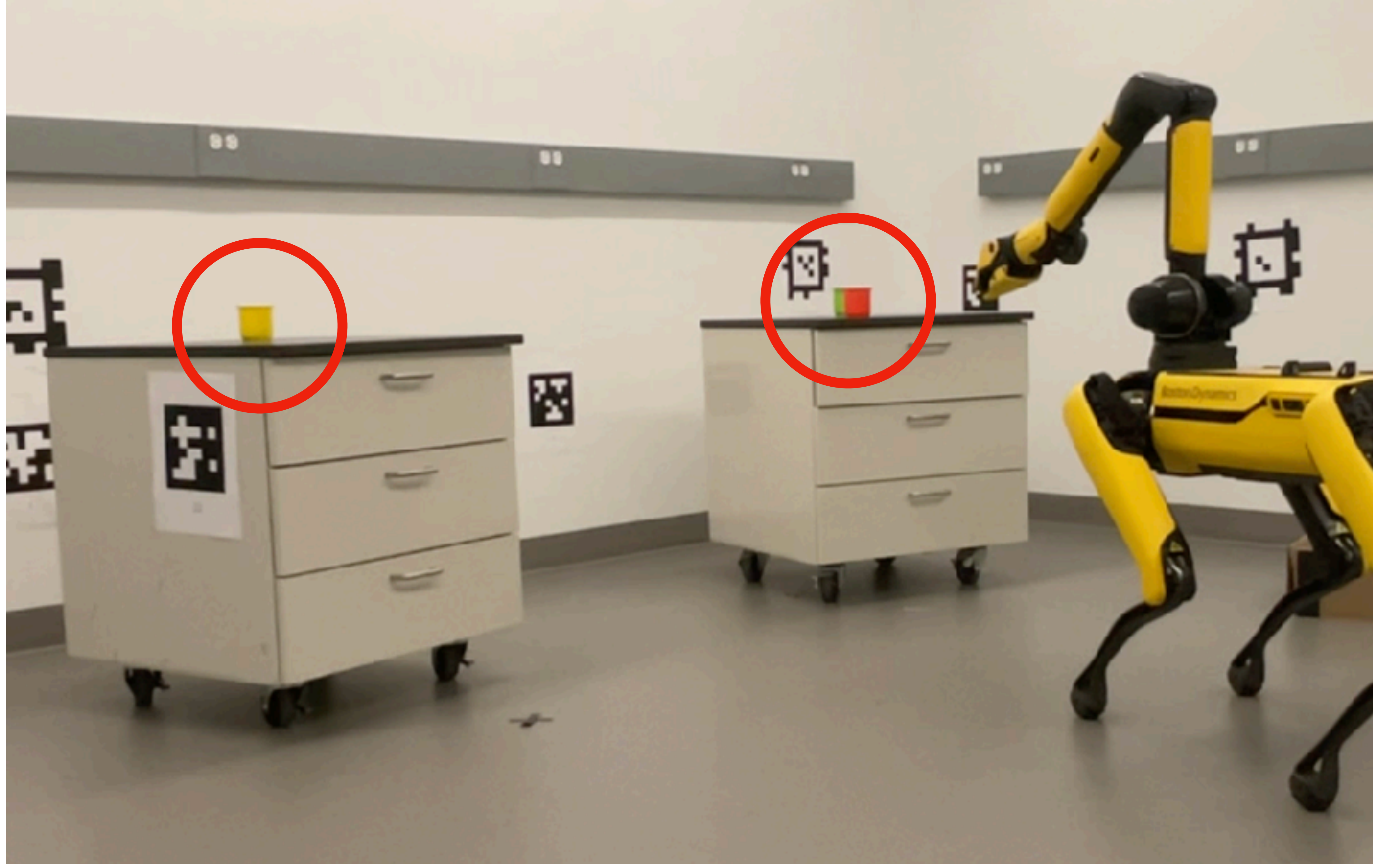


What if we don't know the full state of the world?
For example: "Remove unused objects in a drawer"
We need to represent belief and plan in belief space

Zhao*, McClinton*, ..., Wong†. *"Planning to Perceive: Toward Mobile Manipulation Under Uncertainty In Open-World Environments"*. In Preparation.

Motivating Example

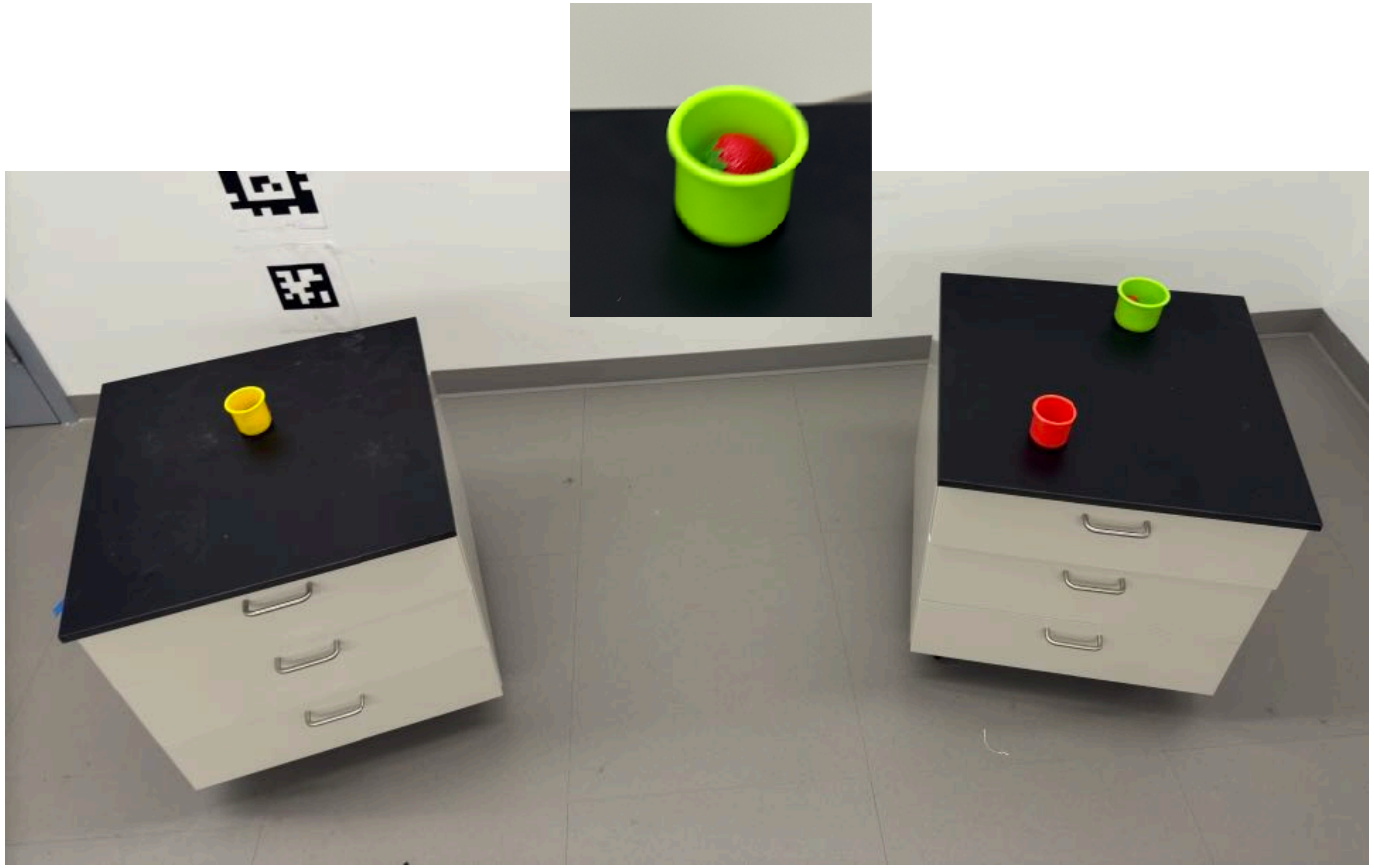
Information Gathering is Needed for Building World State



- We want the robot to operate in open-world environment without full prior knowledge.
  - E.g., "remove empty cups"
- The robot doesn't know some object properties.
  - E.g., whether cups are empty or not.

Motivating Example

Integrating Belief-space Planning for Building World State



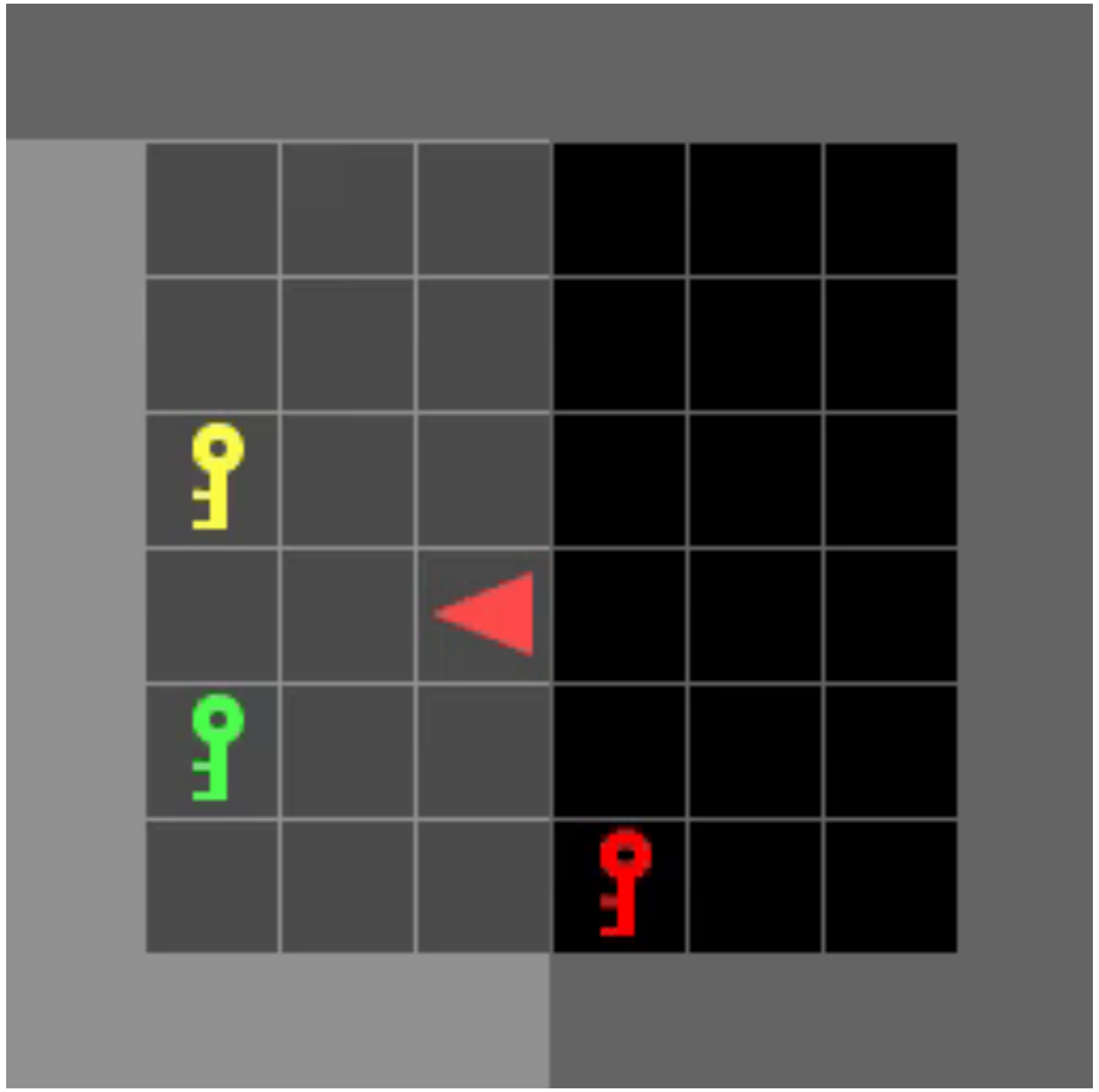
Can robot plan to gather such information? Yes!

It needs:

- Represent its uncertainty on the unknown properties via belief state
- Plan in belief space to take actions to minimize uncertainty

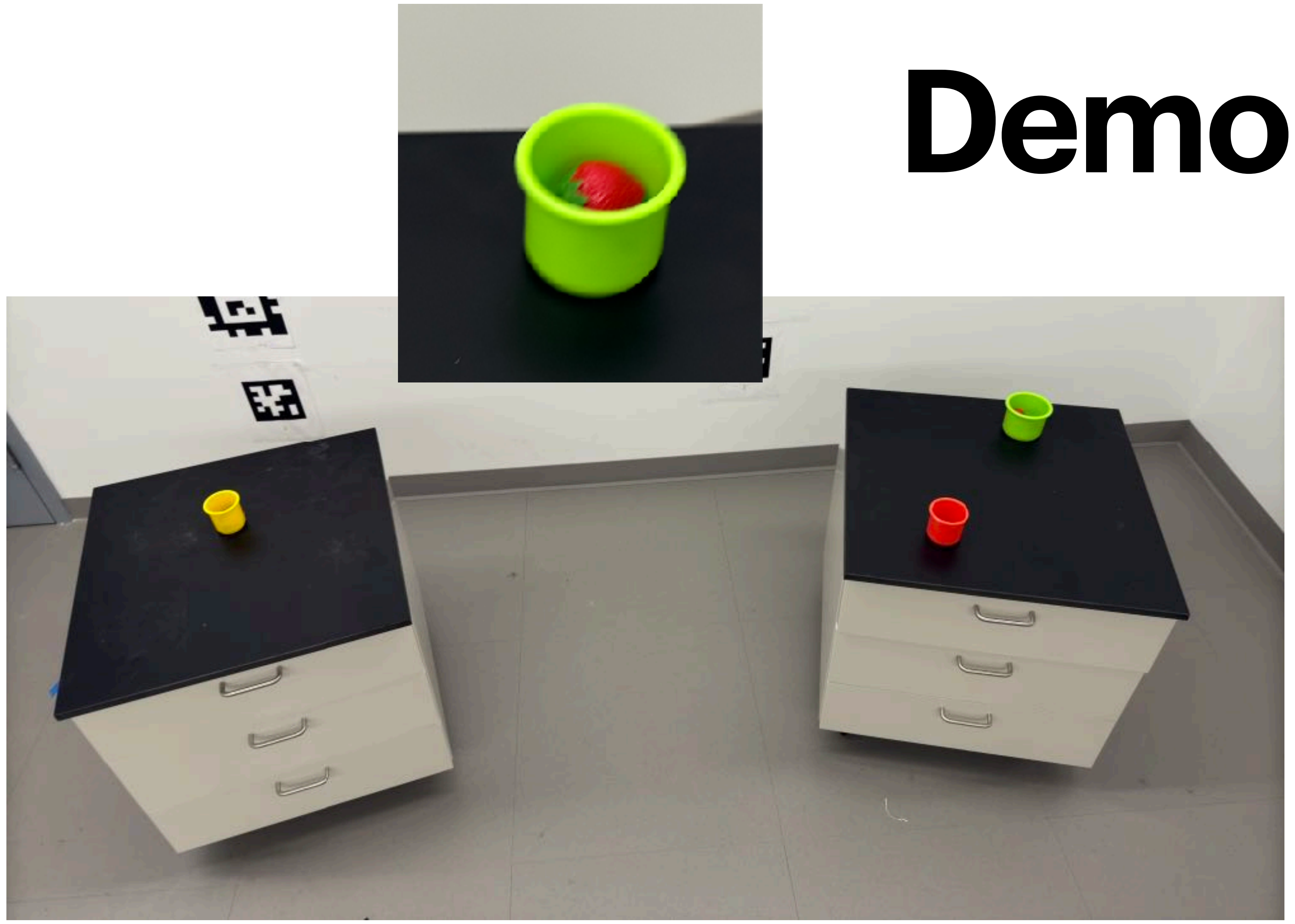
High-level Idea

Integrating Perception and Belief-space Planning Guided by VLMs



- Instruction: “pick up keys in red”
- The agent is deployed in a new room without prior knowledge of:
  - Object Existence
  - Symbolic Goals
  - Object Properties
    (e.g., colors, emptiness)

Demo Task Setup



- Spot robot in a new room without prior knowledge of the world
- It has belief-space operators+skills
- Spot needs to:
  - 1) “See” the objects
  - 2) Perceive the object properties
  - 3) Ground instruction into info gathering subgoals
  - 4) Plan in belief space

Approach: Planning to Perceive

VLM-guided Perception, Grounding, and Belief-space Planning

- Grounding a high-level instruction into *symbolic goals and objects*
  - 1) Using VLM to propose objects robot sees
  - 2) Using VLM to parse the scenes into (a sequence of) subgoals
- Perceiving object relational properties (belief-space predicates)
  - Using VLM to perceive object properties (True/False) and uncertainty
- Belief-space Task Planning
  - Planning with belief-space operators using ternary predicates

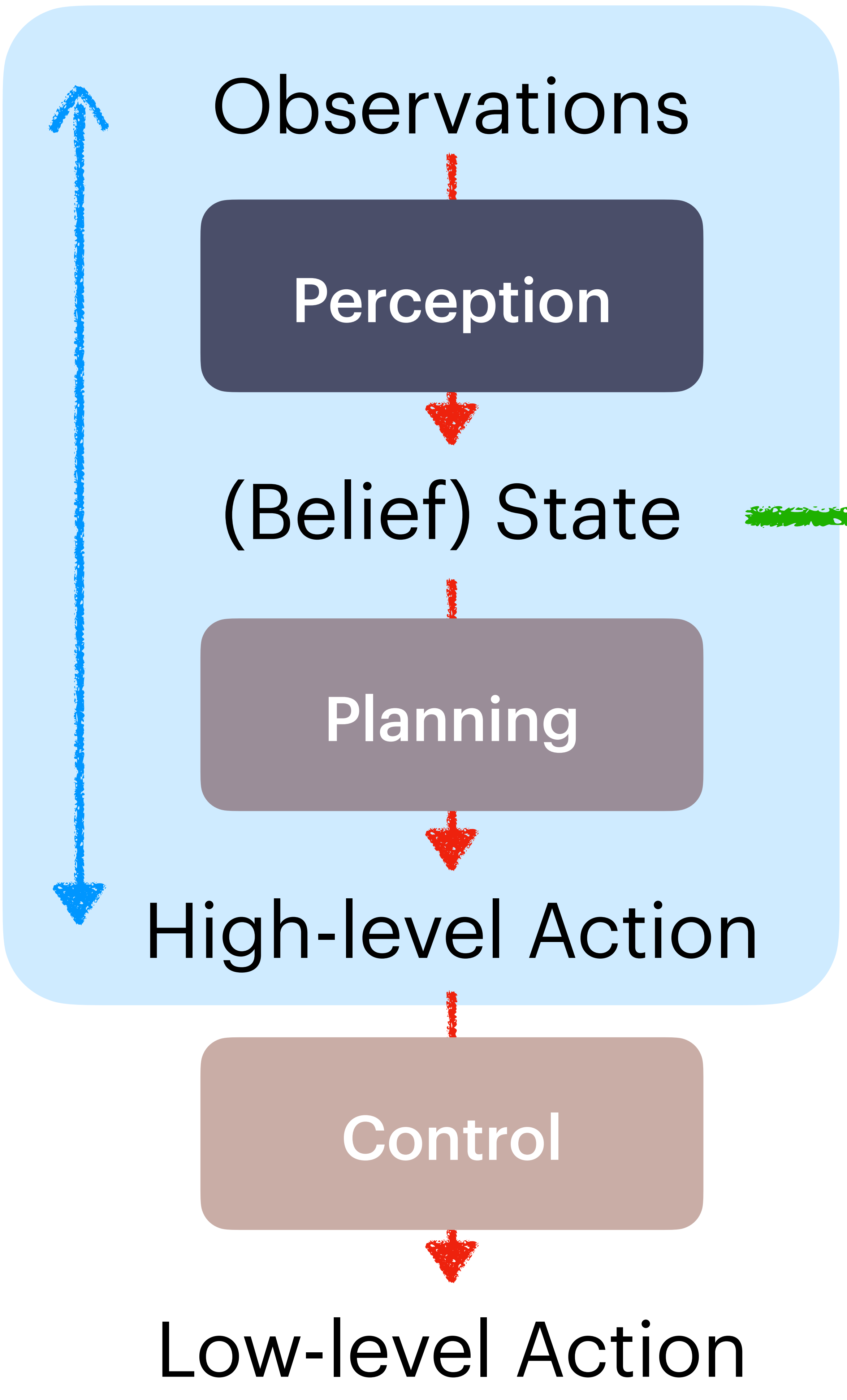
4x

Takeaways

Integrating Perception and Belief-space Planning Guided by VLMs

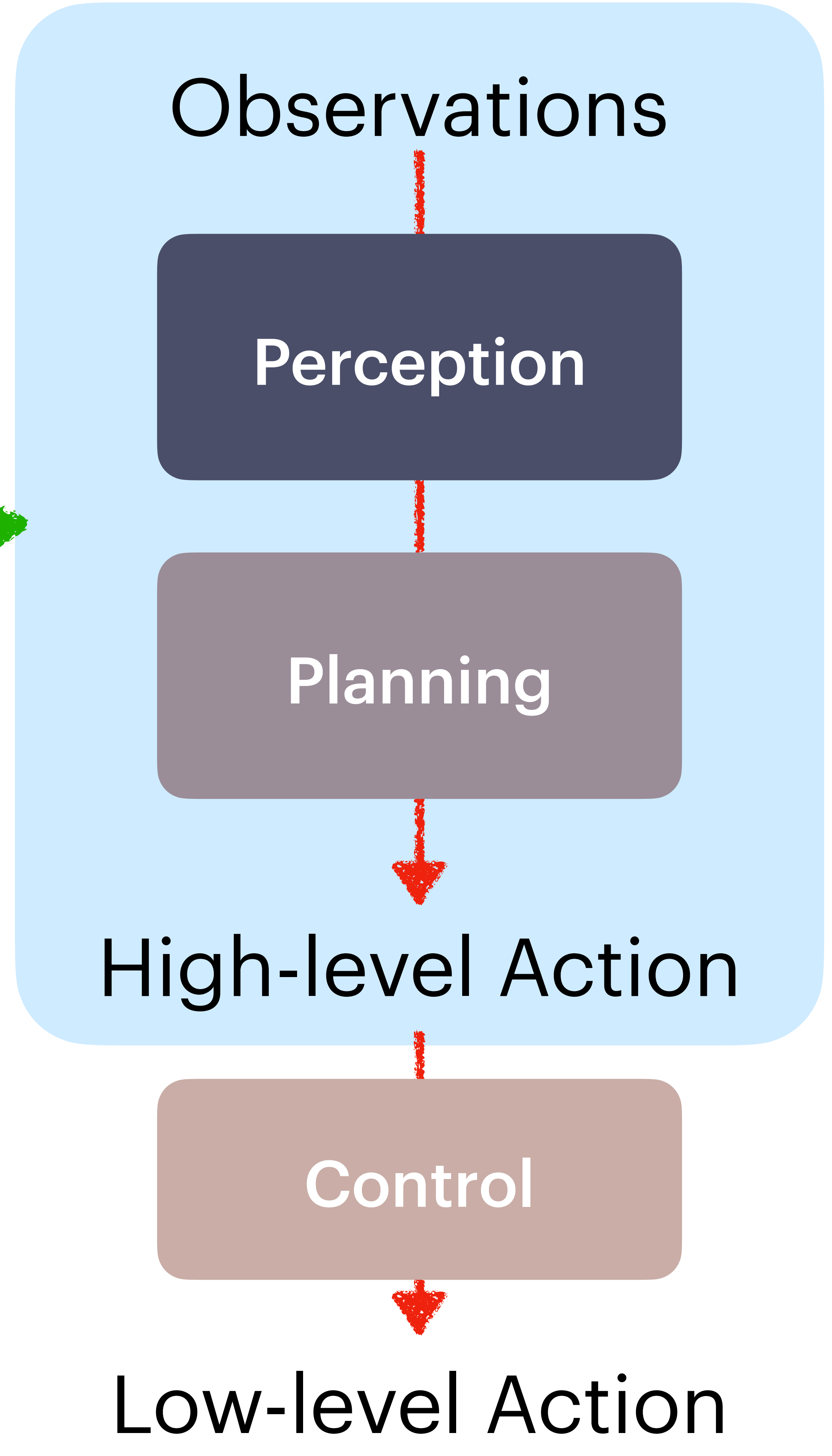
- ✅ Integration of perception and belief-space planning at “inference time” enables strategic information gathering for modeling world state
- ✅ Planning in symbolic belief space enables long-horizon behaviors
- ✅ With a structured pipeline, large models guide the robot to adapt to open-world environment
- ⚠️ The system still needs to hand design belief state representation

Next: Eliminate Explicit State Representation?

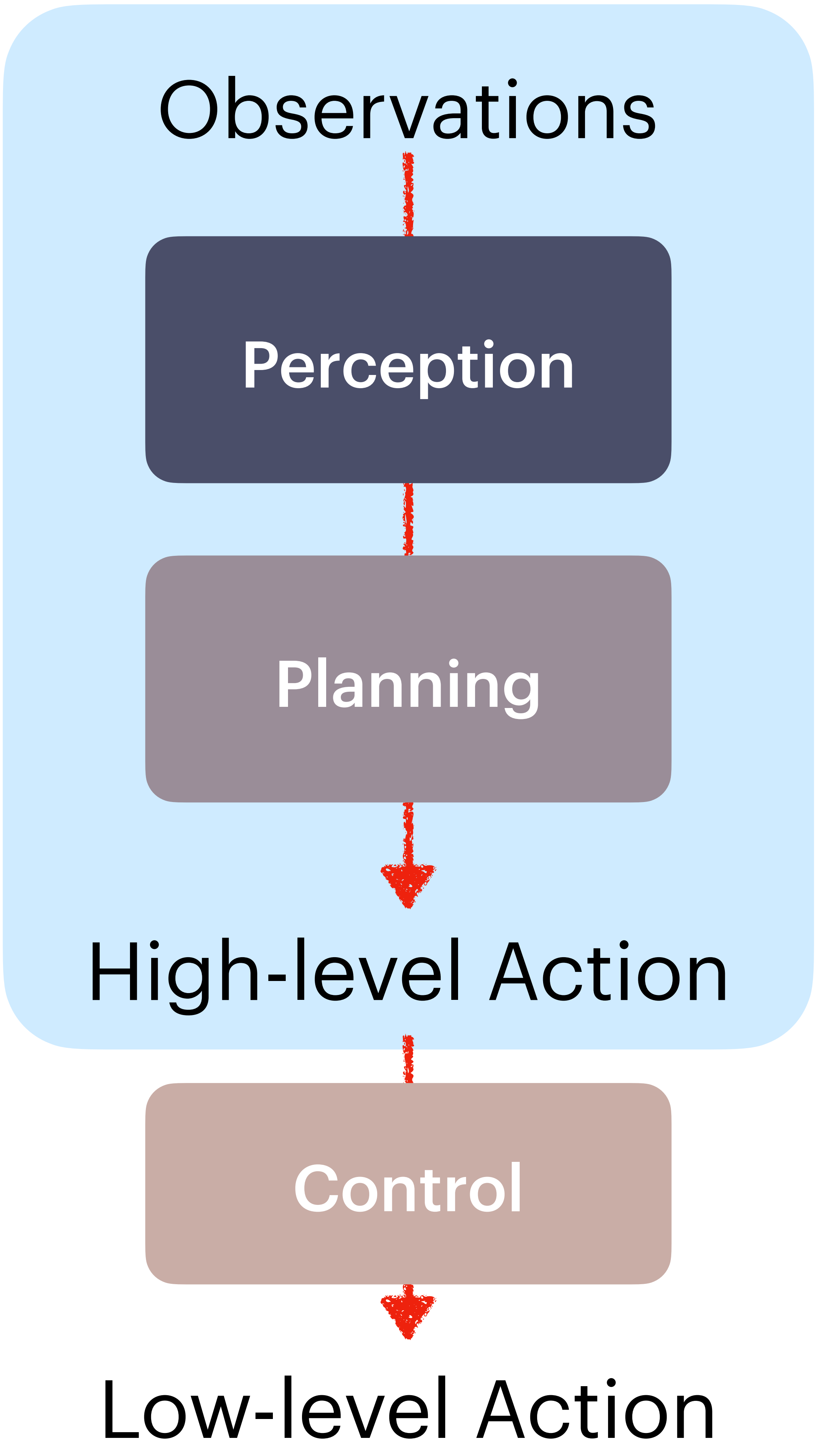


⚠️ It can be very challenging on engineering *explicit representation*: open-world scenarios, diverse objects, uncertainty, partial observability...

💡 Pretrained large models "understand" language and images.
Are they good enough for predicting high-level plan using raw observations?



How Joint World Modeling and Planning?



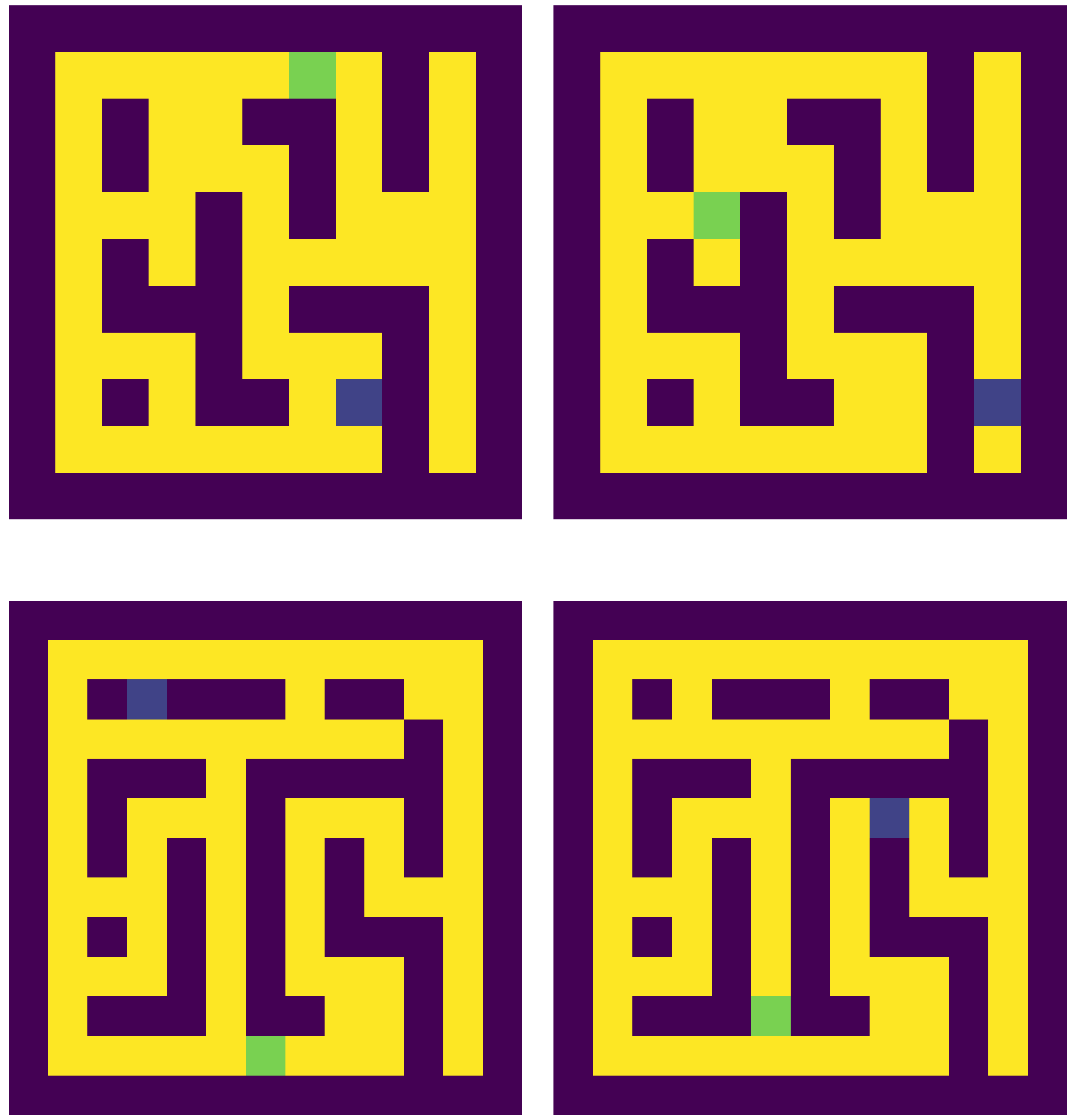
VLM planning
(E.g., Driess et al. 2023)

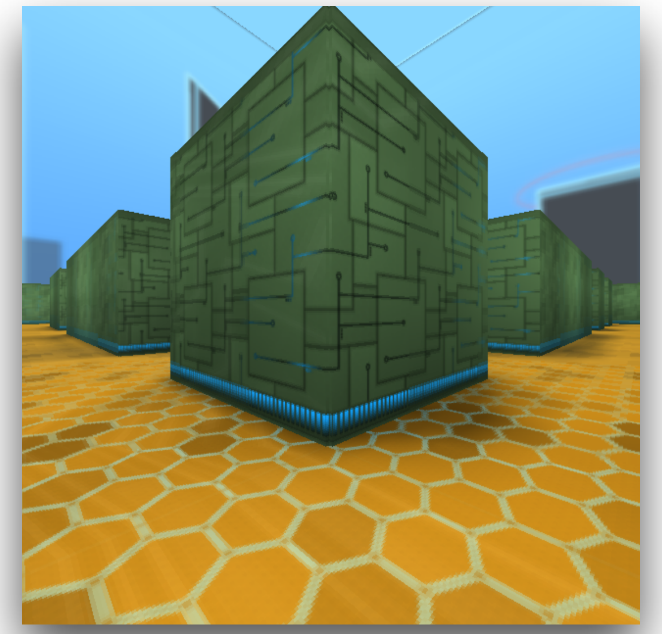
High-level action is a boundary:

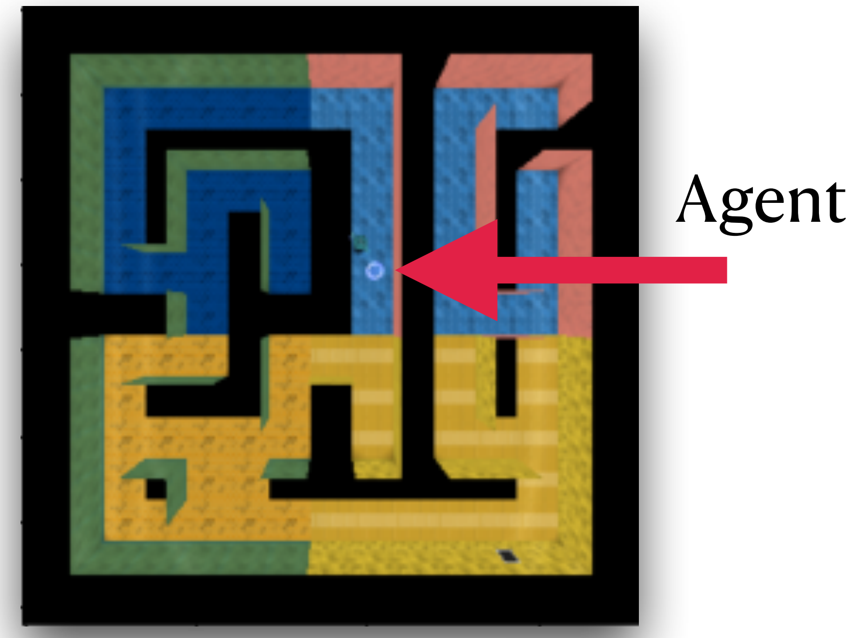
Above: Images to High-level Plan: Data may be share across embodiments

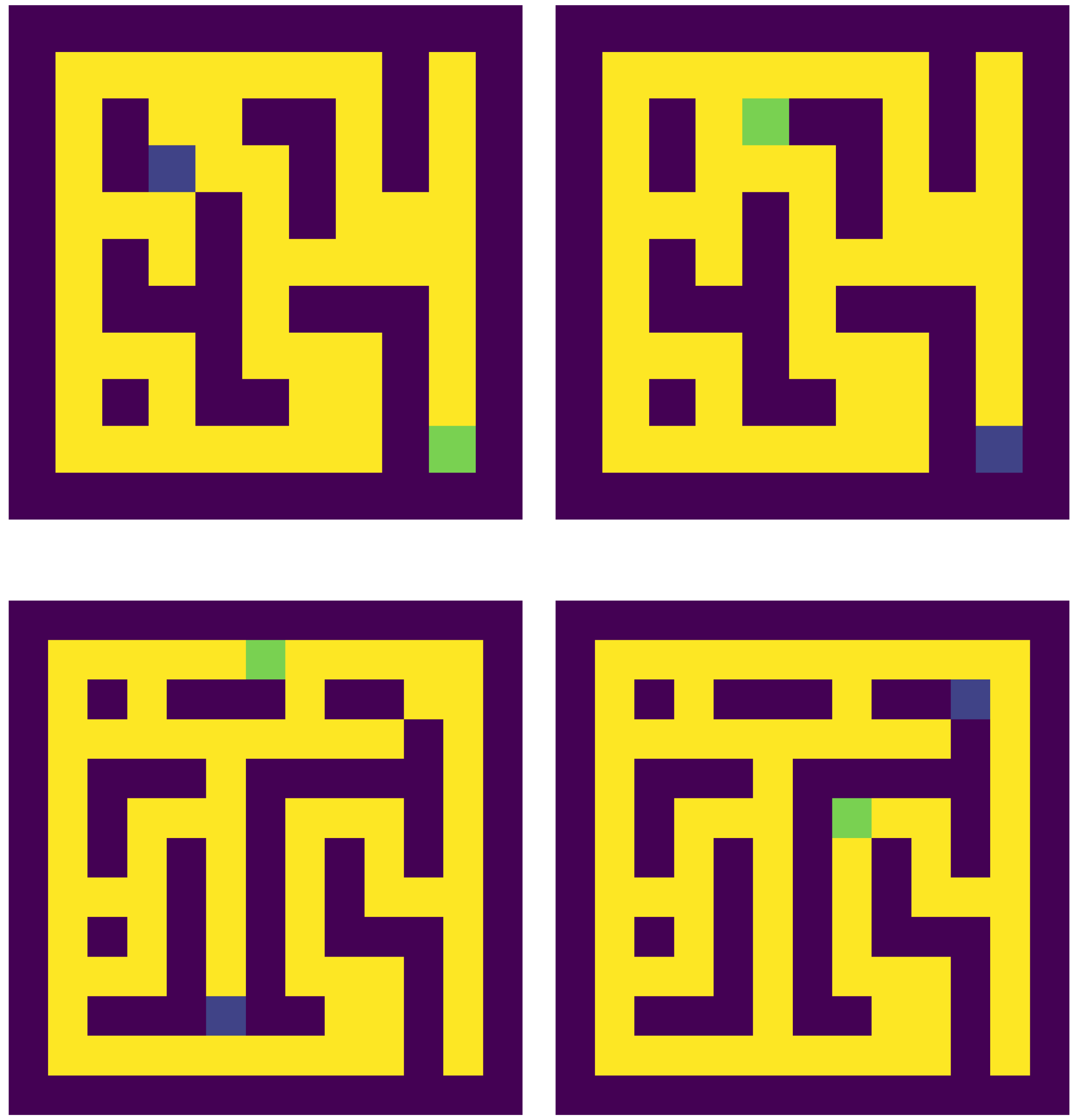
Below: *High-level to Low-level Actions* Physical data is needed and is hard to transfer between robots

Abstract High-level Maps for Guided Planning









Learning Navigation Policies

Given State, Map and Goal
Plan the actions:

$$\text{Plan}\left(s, m, g, f_\phi\right)$$

Training

Zero-shot Navigation

Zhao, Wong. *"Learning to Navigate in Mazes with Novel Layouts Using Abstract Top-down Maps"*. RLC 2024.

Summary

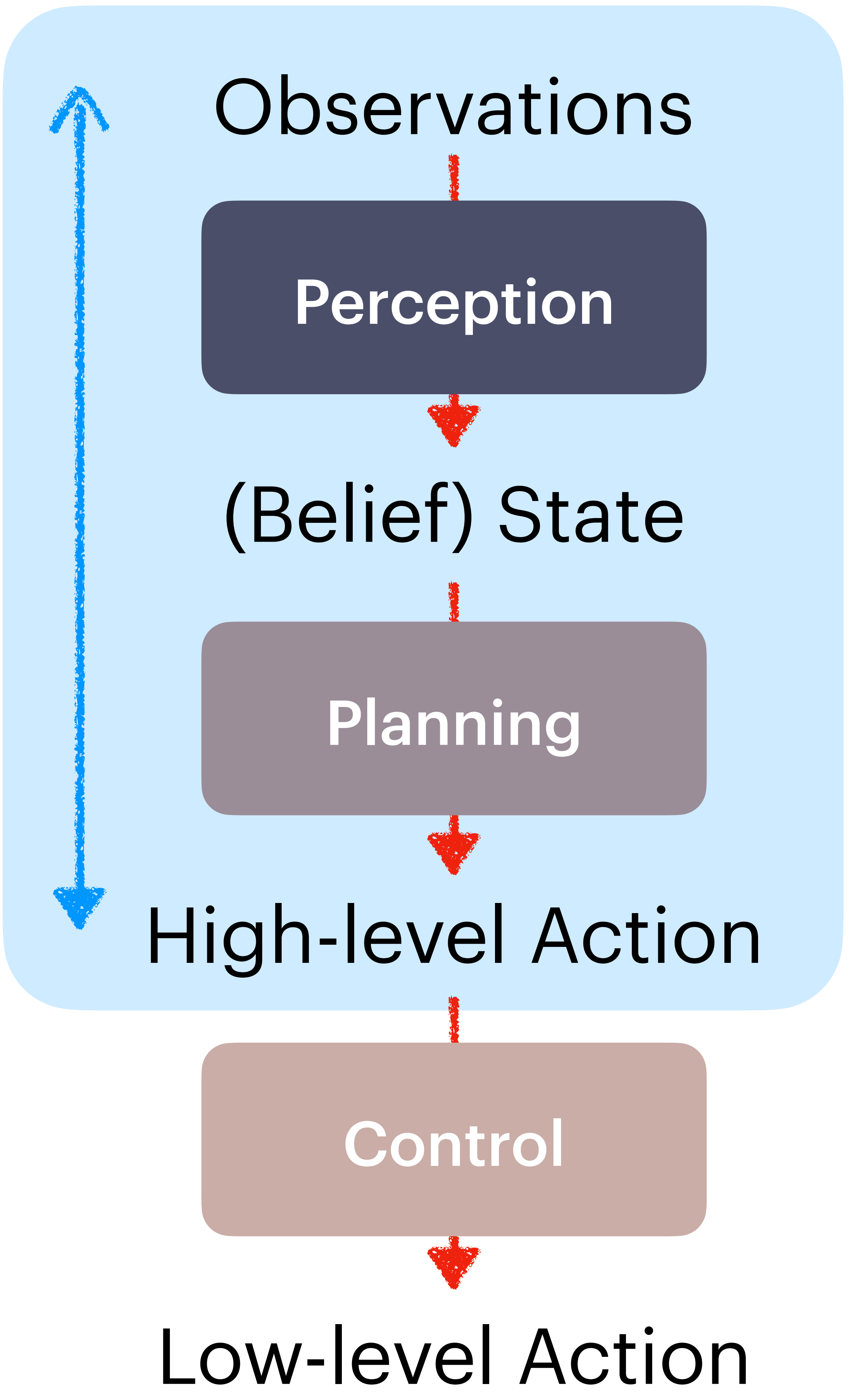
Lossy Abstraction of World Representation and Planning

- Lossy abstraction enables computationally feasible high-level planning in complex environments, but effectively grounding actions remains a core challenge.
- Instead of end-to-end learning, structured approaches that integrate planning with perception and control provide more efficient solution

Takeaways

Summary

- Effective planning is crucial for agents to perform long-horizon and challenging tasks.
- Integration of structured learning requires much less data
- Balancing lossless and lossy abstraction is needed
    - For lossy representation, rounding high-level plans into low-level actions is a challenge that needs effective solutions
- Progressing from separate perception and planning modules to integrated systems is challenging but could provide more flexibility



Thank you!

Linfeng Zhao

My research:

Using Learning for World Modeling and Planning with Structured Approaches

Microsoft Research

Northeastern University

Northeastern University Khoury College of Computer Sciences

Boston Dynamics AI Institute

amazon | science

Meta