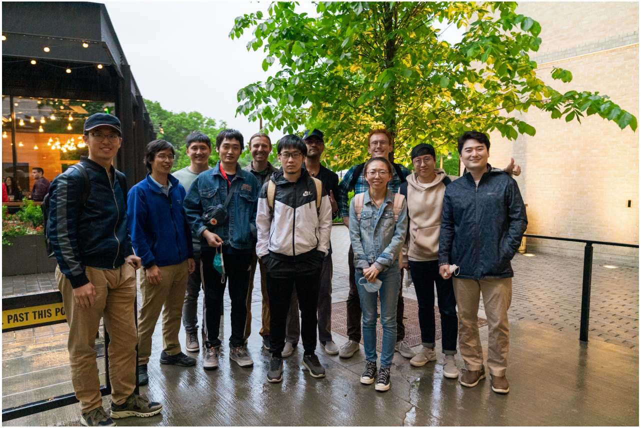
Symmetry in Decision Making

Lawson L.S. Wong

lsw@ccs.neu.edu
Khoury College of Computer Sciences
Northeastern University

November 3, 2023

Northeastern University
Khoury College of
Computer Sciences



GRAIL

Generalizable Robotics and
Artificial Intelligence Laboratory

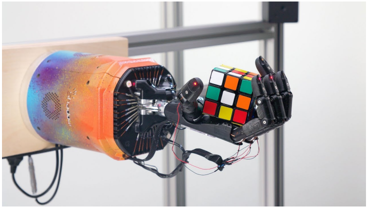
Reinforcement learning is a promising framework for generalizable robotics.

Excitement

[Mnih et al. 2015]

[Silver et al. 2016]

[Levine et al. 2016]



[OpenAI 2019]

Excitement

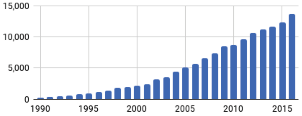


Figure 1: Growth of published reinforcement learning papers. Shown are the number of RL-related publications (y-axis) per year (x-axis) scraped from Google Scholar searches.

[Henderson et al. 2018]

Reinforcement Learning Papers Accepted to ICLR 2020

reinforcement-learning

Dec 24, 2019 • Seungjae Ryan Lee

Show 10 entries

Search:

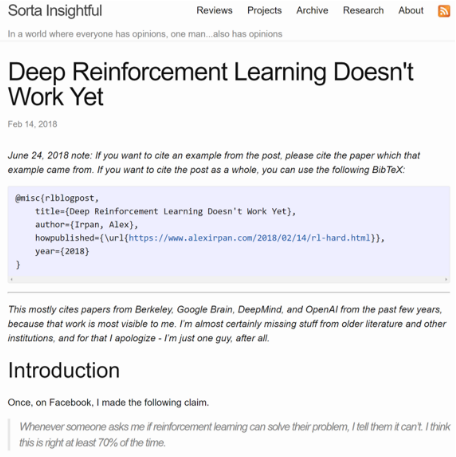
Rank	Average Rating	Title	Ratings	Variance	Decision
1	8.00	Dynamics-aware Unsupervised Skill Discovery	8 8 8	0.00	Accept (Talk)
1	8.00	Contrastive Learning Of Structured World Models	8 8 8	0.00	Accept (Talk)
1	8.00	Implementation Matters In Deep Rl: A Case Study On Ppo And Trpo	8 8 8	0.00	Accept (Talk)
1	8.00	Gendice: Generalized Offline Estimation Of Stationary Values	8 8 8	0.00	Accept (Talk)
1	8.00	Causal Discovery With Reinforcement Learning	8 8 8	0.00	Accept (Talk)
2	7.33	Is A Good Representation Sufficient For Sample Efficient Reinforcement Learning?	8 8 6	0.89	Accept (Spotlight)
2	7.33	Harnessing Structures For Value-based Planning And Reinforcement Learning	6 8 8	0.89	Accept (Talk)
2	7.33	Explain Your Move: Understanding Agent Actions Using Focused Feature Saliency	6 8 8	0.89	Accept (Poster)
2	7.33	Meta-q-learning	8 8 6	0.89	Accept (Talk)
2	7.33	Discriminative Particle Filter Reinforcement Learning For Complex Partial Observations	8 6 8	0.89	Accept (Poster)

Showing 1 to 10 of 106 entries

Previous 1 2 3 4 5 ... 11 Next

`https://www.endtoend.ai/blog/iclr2020-rl/`

Are we done?


Sorta Insightful

Reviews Projects Archive Research About

In a world where everyone has opinions, one man...also has opinions

Deep Reinforcement Learning Doesn't Work Yet

Feb 14, 2018

June 24, 2018 note: If you want to cite an example from the post, please cite the paper which that example came from. If you want to cite the post as a whole, you can use the following BibTeX:

```
@misc{rlblogpost,
    title={Deep Reinforcement Learning Doesn't Work Yet},
    author={Irpan, Alex},
    howpublished={\url{https://www.alexirpan.com/2018/02/14/rl-hard.html}},
    year={2018}
}
```

This mostly cites papers from Berkeley, Google Brain, DeepMind, and OpenAI from the past few years, because that work is most visible to me. I'm almost certainly missing stuff from older literature and other institutions, and for that I apologize - I'm just one guy, after all.

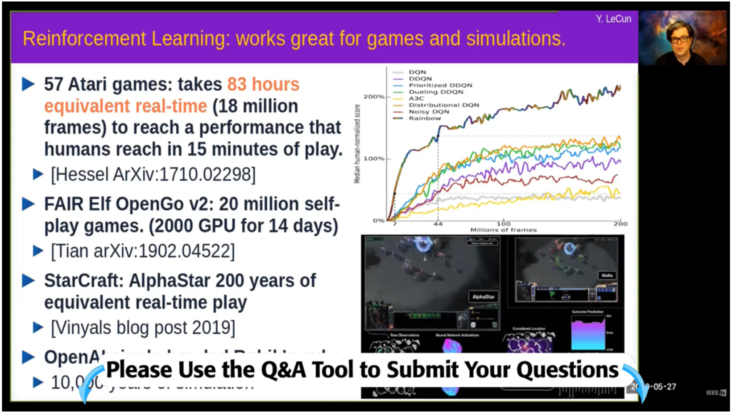
Introduction

Once, on Facebook, I made the following claim.

> Whenever someone asks me if reinforcement learning can solve their problem, I tell them it can't. I think this is right at least 70% of the time.

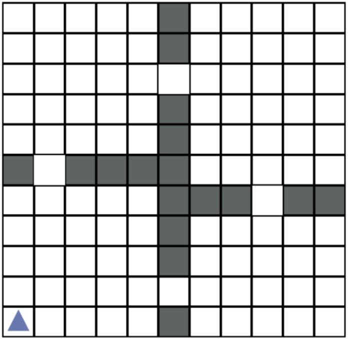
`https://www.alexirpan.com/2018/02/14/rl-hard.html`

Problem: Sample complexity / learning speed



https://ieeetv.ieee.org/conference-highlights/
self-supervised-learning-world-models-icra-2020
(See 11m00s in video)

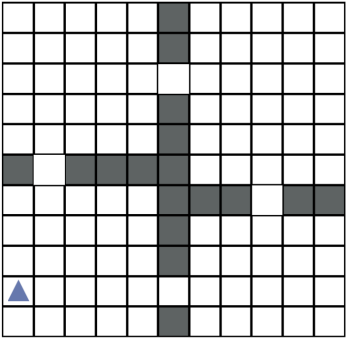
What is reinforcement learning?



(a) The Four Rooms Domain

State: $(0, 0)$

What is reinforcement learning?



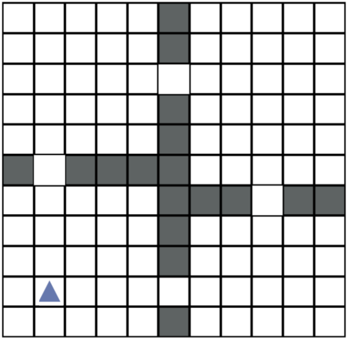
(a) The Four Rooms Domain

Action: up

State: $(0, 1)$

Reward: 0

What is reinforcement learning?



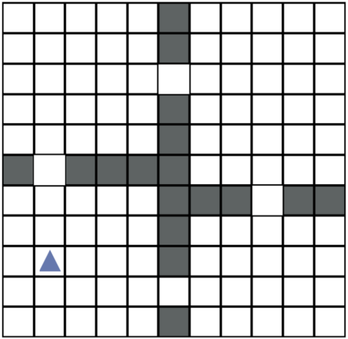
(a) The Four Rooms Domain

Action: up

State: $(1, 1)$

Reward: 0

What is reinforcement learning?



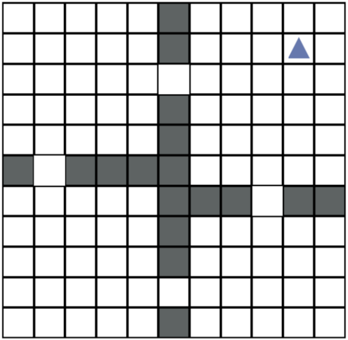
(a) The Four Rooms Domain

Action: up
State: $(1, 2)$
Reward: 0

What is reinforcement learning?

(many time steps later ...)

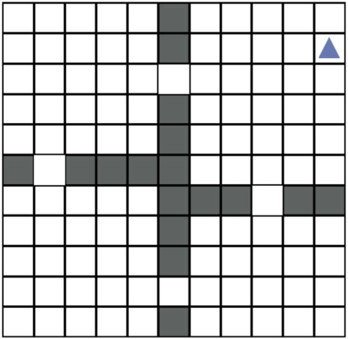
What is reinforcement learning?



(a) The Four Rooms Domain

Action: `right`
State: $(9, 9)$
Reward: 0

What is reinforcement learning?



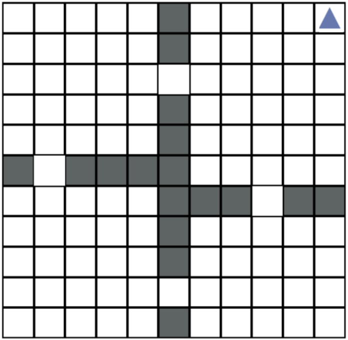
(a) The Four Rooms Domain

Action: `right`

State: $(9, 10)$

Reward: 0

What is reinforcement learning?



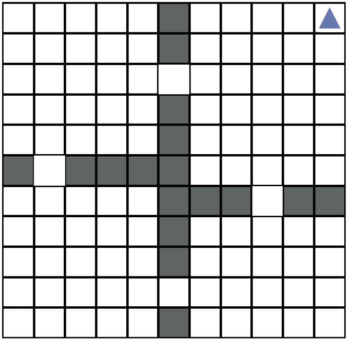
(a) The Four Rooms Domain

Action: up
State: $(10, 10)$
Reward: $+1$

What is reinforcement learning?

(reset, try again)

What is reinforcement learning?

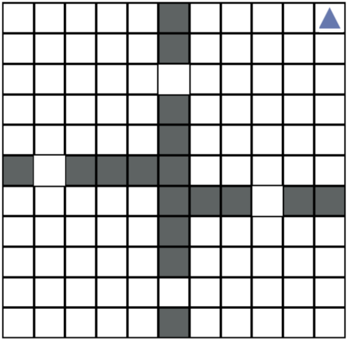


(a) The Four Rooms Domain

Action: up
State: $(10, 10)$
Reward: $+1$

Objective: Find **policy** (state-action mapping)

What is reinforcement learning?

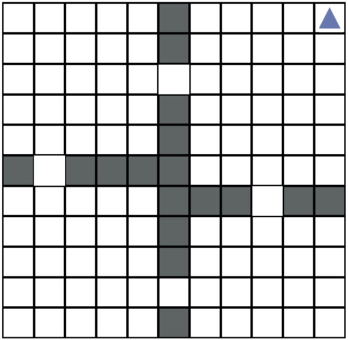


(a) The Four Rooms Domain

Action: up
State: $(10, 10)$
Reward: $+1$

Objective: Find **policy** (state-action mapping) that maximizes **expected sum of rewards**

What is reinforcement learning?



(a) The Four Rooms Domain

Action: up
State: $(10, 10)$
Reward: $+1$

Objective: Find **policy** (state-action mapping) that maximizes **expected sum of rewards**

Key issues:

- How good is $+1$?
- Credit assignment
- Sparse reward
- Exploration vs. exploitation

Thesis

Reinforcement learning is very general

Thesis

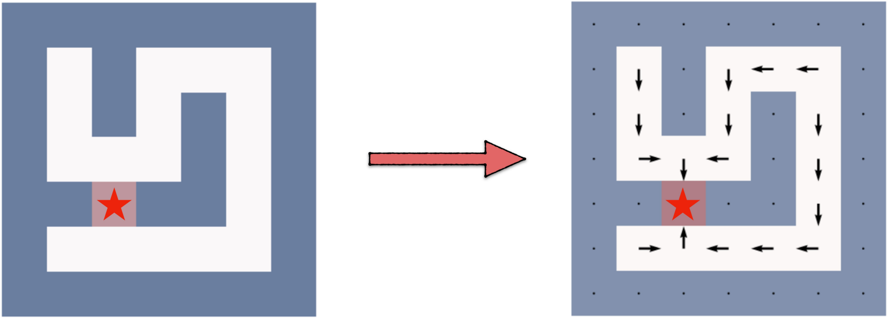
Reinforcement learning is very general ; but reinforcement learning algorithms are slow **because** they are too general.

What additional guidance can we give RL agents so they are more specific but **useful**?

Symmetry in RL and decision making

- **Exploitation**
  Integrating symmetry into differentiable planning
- Opportunity
  Robot navigation with coarse maps
- Exploration
  Speculative wish list

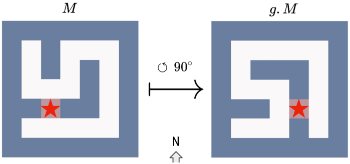
Symmetry in 2-D path-planning



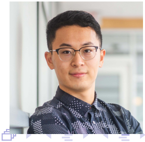
Find shortest path / optimal actions to the goal location (red)

Symmetry in 2-D path-planning

What does the symmetry look like?



Linfeng Zhao *et al.* Integrating symmetry into differentiable planning with steerable convolutions. ICLR 2023.



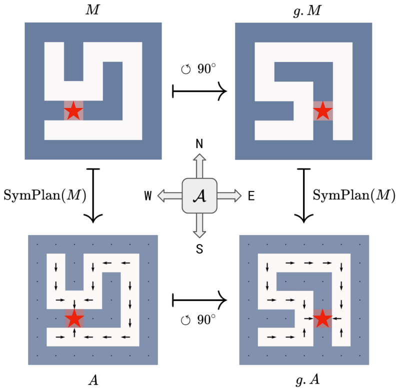
Symmetry in 2-D path-planning

What does the symmetry look like?

They can be described by

Equivariance

$$\circlearrowleft 90^\circ \circ (\mathit{Plan}(M)) = \mathit{Plan}(\circlearrowleft 90^\circ \circ M)$$



Linfeng Zhao *et al.* Integrating symmetry into differentiable planning with steerable convolutions. ICLR 2023.

Symmetric planning: Background

- Classical planning:
  Fox & Long, Pochter *et al.*, Domshlak *et al.*, Shleyfman *et al.*, Sievers *et al.*, . . .
  - Perform orbit search

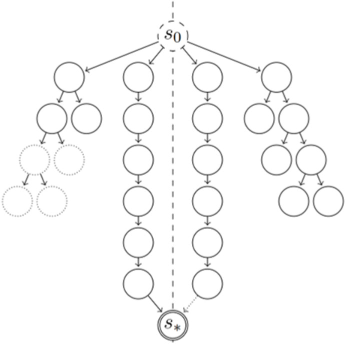


Figure from Shleyfman *et al.*

Symmetric planning: Background

- Classical planning:
  Fox & Long, Pochter *et al.*, Domshlak *et al.*, Shleyfman *et al.*, Sievers *et al.*, ...
  - Perform orbit search
- Symmetries in MDPs:
  Ravindran & Barto, Ravindran, Ferns *et al.*, Narayanamurthy & Ravindran, van der Pol *et al.*, ...
  - Construct "quotient MDP"

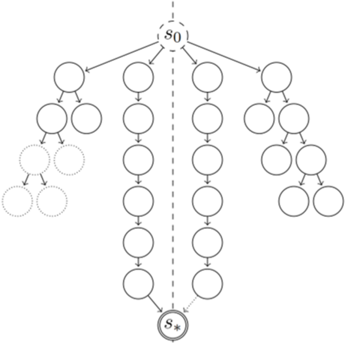


Figure from Shleyfman *et al.*

Symmetric planning: Background

Differentiable value iteration / planning algorithms:
Make planning algorithms fully differentiable

- Value iteration: Tamar *et al.*, Niu *et al.*, Lee *et al.*, Xu *et al.*, Chaplot *et al.*, Deac *et al.*, ...
- Differentiable planning: Oh *et al.*, Karkus *et al.*, Weber *et al.*, Srinivas *et al.*, Schrittwieser *et al.*, Amos & Yarats, Wang & Ba, Guez *et al.*, Hafner *et al.*, Pong *et al.*, Clavera *et al.*, Hansen *et al.*, ...

Bellman equation

$$Q(s,a) = R(s,a) + \gamma \sum_{s'} \mathbb{P}(s'|s,a') V(s')$$

$$V(s) = \max_a Q(s,a)$$

Learn R_a, P_a

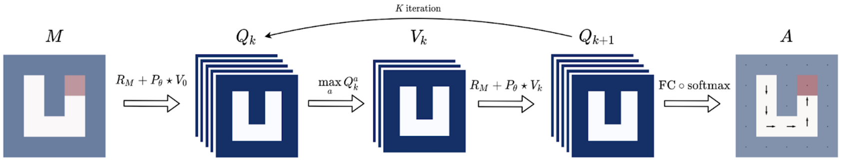
$$Q_a = R_a + \mathrm{Conv2D}(V, P_a)$$

$$V = \max_a Q_a$$

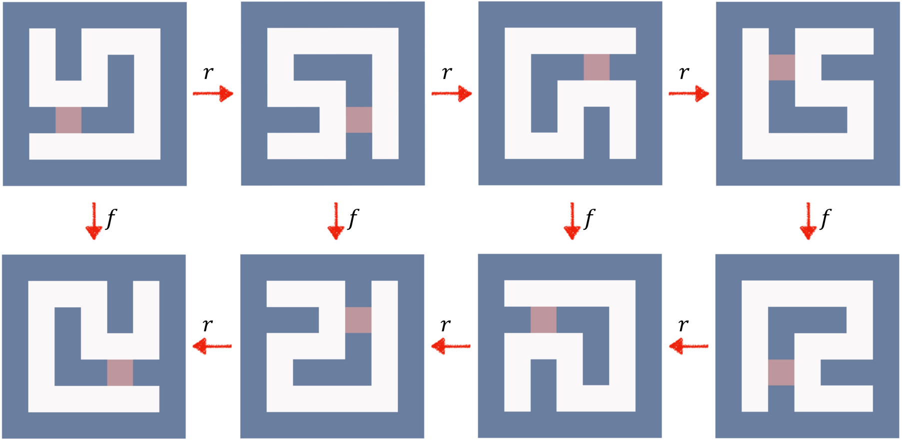
Symmetric planning: Background

Differentiable value iteration / planning algorithms:
Make planning algorithms fully differentiable

- Value iteration: Tamar *et al.*, Niu *et al.*, Lee *et al.*, Xu *et al.*, Chaplot *et al.*, Deac *et al.*, ...
- Differentiable planning: Oh *et al.*, Karkus *et al.*, Weber *et al.*, Srinivas *et al.*, Schrittwieser *et al.*, Amos & Yarats, Wang & Ba, Guez *et al.*, Hafner *et al.*, Pong *et al.*, Clavera *et al.*, Hansen *et al.*, ...



Symmetry in 2-D path-planning: D_4



Exploiting symmetry in 2-D path-planning

Every update is equivariant — Local Equivariance

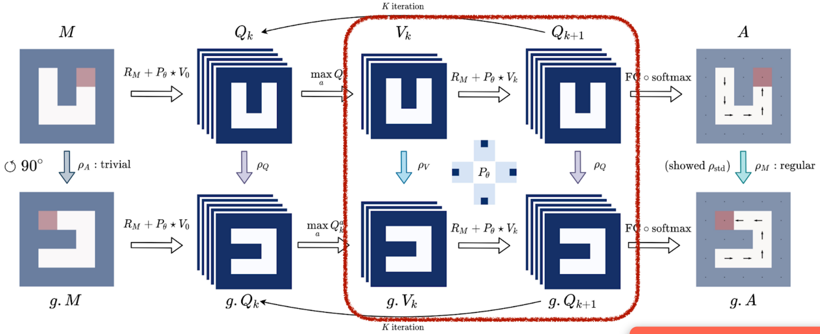
Entire planning is equivariant — Global Equivariance

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

- Use steerable convolution, equivariant to rotation and reflection:

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

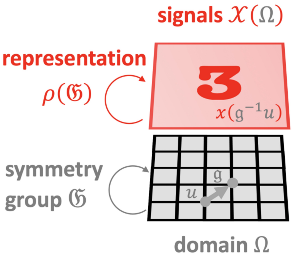
Symmetric planning: SymVIN

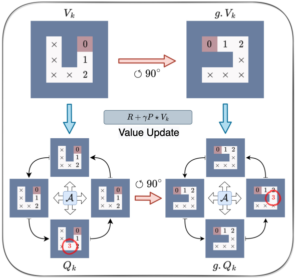


Every pair is equivariant

We use steerable convolutions to integrate symmetry in VINs.

Symmetric planning: Insights



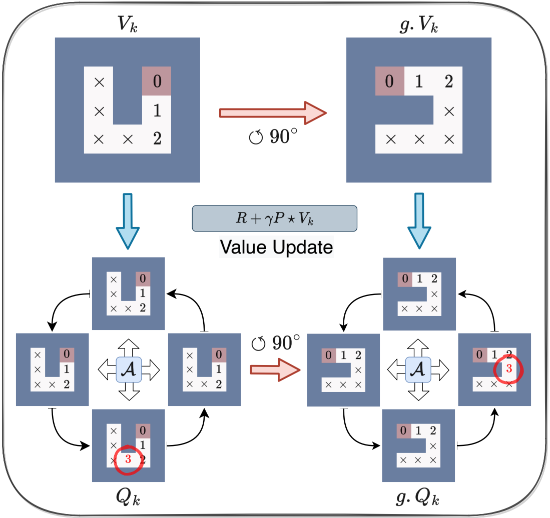


1: Represent (value) functions as "fields"

2: Value iteration as convolution (network)

Bronstein et al. (2021): Geometric Deep Learning: Grids, Groups, Graphs, Geodesics, and Gauges. arXiv.

Symmetric planning: Bellman back-up



Symmetric planning: Theory

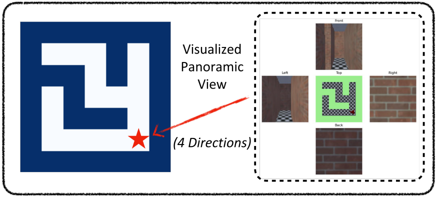
Theorem 1 (informal): **Value iteration** for **path planning*** is equivariant to translation, rotation, and reflection

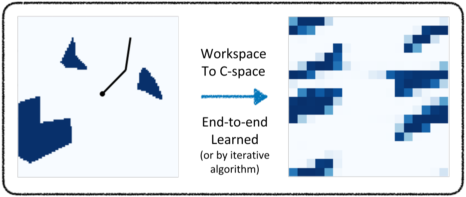
Theorem 2 (informal): **Value iteration** for **path planning*** is a form of **steerable convolution network****

*: 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

Symmetric planning: Domains



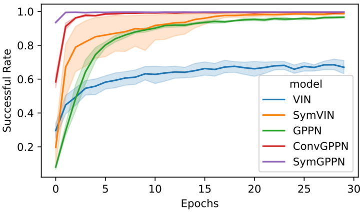


2D and Visual
Maze Navigation

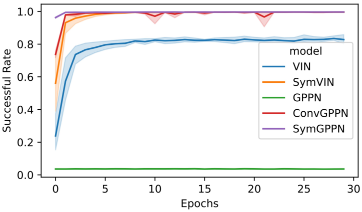
2-DOF Manipulation
In Workspace and C-space

Symmetric planning: Efficient training

2D Maze Navigation



2-DOF Manipulation



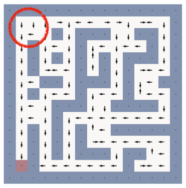
- Training curves on 15x15 maps
- All networks use 3x3 filters, fixed 30 iterations (shared convolution layers)
- SymGPPN — recurrent version of VIN with steerable convolution

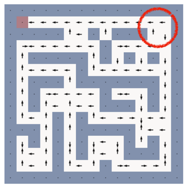
Symmetric planning: Higher success rates

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

Symmetric planning: Equivariance

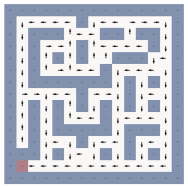
Feed in M and $\circlearrowleft 90^\circ \circ M$

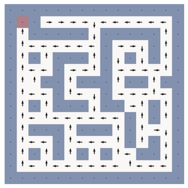




VIN output doesn't satisfy equivariance

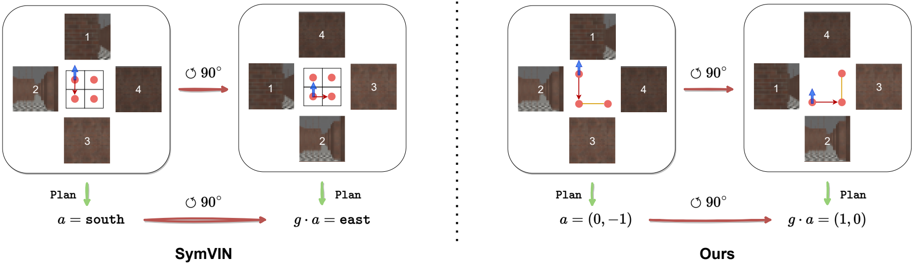
Feed in M and $\circlearrowleft 90^\circ \circ M$





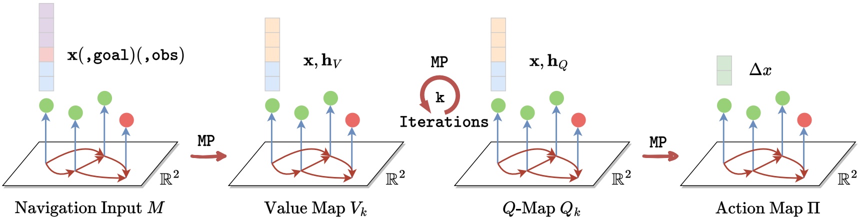
SymVIN guarantees output is equivariant

Extension: Symmetric planning on graphs



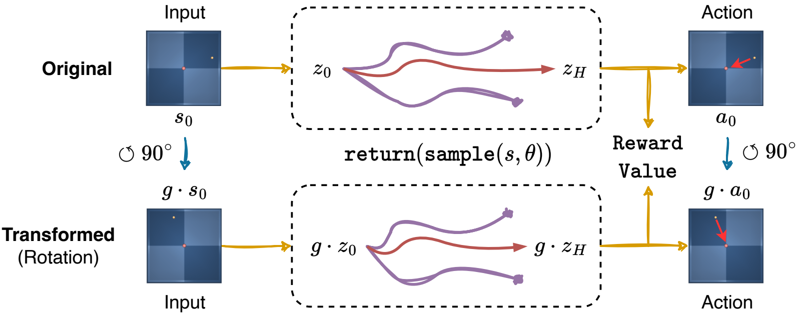
Linfeng Zhao*, Hongyu Li* *et al.* $E(2)$-equivariant graph planning for navigation. In submission.

Extension: Symmetric planning on graphs



Extension: Symmetric planning on graphs

Extension: Sampling-based symmetric planning



Linfeng Zhao *et al.* Can Euclidean symmetry help in reinforcement learning and planning? In submission.

See Linfeng's poster!

Symmetry in RL and decision making

- Exploitation
  Integrating symmetry into differentiable planning
- **Opportunity**
  Robot navigation with coarse maps
- Exploration
  Speculative wish list

Navigation

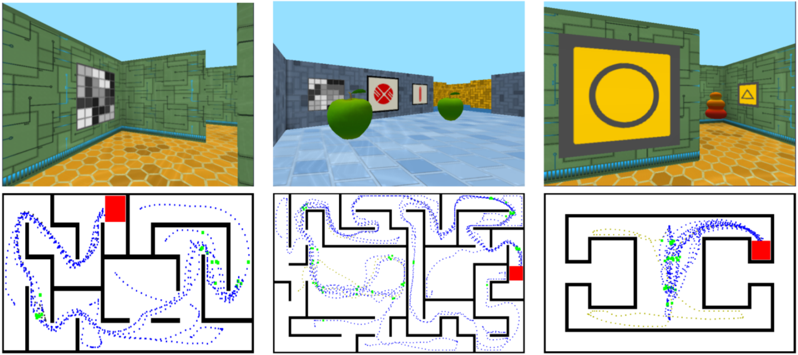
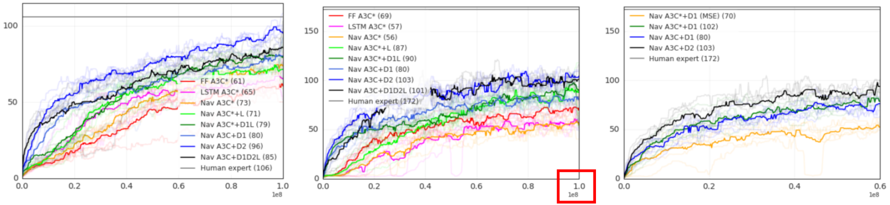


Figure 1: Views from a small 5×10 maze, a large 9×15 maze and an I-maze, with corresponding maze layouts and sample agent trajectories. The mazes, which will be made public, have different textures and visual cues as well as exploration rewards and goals (shown right).

Learning to navigate in complex environments [Mirowski et al. 2017]

Navigation



(d) Random Goal maze (small)

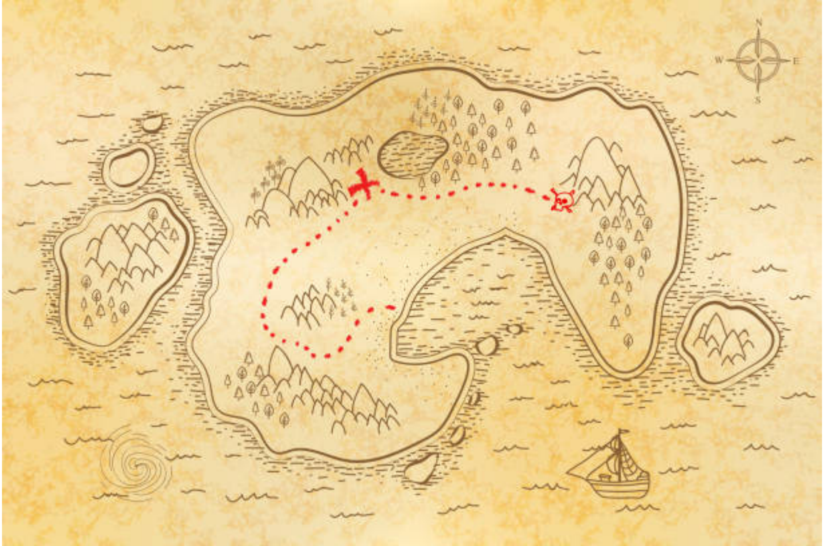
(e) Random Goal maze (large)

(f) Random Goal maze (large): different formulation of depth prediction

Figure 3: Rewards achieved by the agents on 5 different tasks: two static mazes (small and large) with fixed goals, two static mazes with comparable layout but with dynamic goals and the I-maze. Results are averaged over the top 5 random hyperparameters for each agent-task configuration. Star in the label indicates the use of reward clipping. Please see text for more details.

Learning to navigate in complex environments [Mirowski et al. 2017]

Navigation **with a map**



Navigation **with a map**

Whether you're an alumnus or alumna returning to campus, or a prospective student seeing Northeastern for the first time, we'll make every effort to ensure you enjoy your visit. We offer 60-minute campus tours on weekdays year-round, and on Saturdays from September through June. Learn more at northeastern.edu/admissions/connect/visit.

Campus Map

Visiting Northeastern

- P Parking (permit required)
- V Visitor parking (paid)
- Handicapped parking
- Handicapped-accessible entrance
- Bluebikes
- Emergency phone
- One-way street
- Pedestrian walkway
- Tunnel entrance

Academic and Service Buildings

C1 140 The Fenway (140)
E2 177 Huntington (177)
E2 216 Massachusetts Ave (216)
E2 236 Huntington (236)
D2 271 Huntington (271)
B1 O'Bryant African American Institute (AF)
C3 Alumni Center (CP)
B2 Architecture Studio (AS)
D1 Asian American Center (AC)
B2 Barletta Natatorium (BN)
A1 Behrakis Health Sciences Center (BK)
E2 Belvidere Place (BV)
C2 Blackman Auditorium (BL)
C2 Cabot Physical Education Center (CB)
C1 Cahners Hall (CA)
B1 Cargill Hall (CG)
D2 Catholic Center (CC)
B2 Churchill Hall (CH)
C3 Columbus Place (CP)
C2 Cullinane Hall (CN)
C2 Curry Student Center (CSC)
C1 Cushing Hall (CU)
B2 Dana Research Center (DA)
B1 Dockser Hall (DK)
C2 Dodge Hall (DG)
C2 East Village (EV)
B2 Egan Research Center (EC)
C2 Ell Hall (EL)
D2 Fenway Center (FC)
B2 Forsyth Building (FR)
C2 Hastings Hall (HS)
C2 Hayden Hall (HA)
D2 Hillel-Frager (HF)
B2 Holmes Hall (HO)
C2 Hurtig Hall (HT)
B3 Interdisciplinary Science and Engineering Complex (ISEC)
A2 International Village (IV)
B1 Kariotis Hall (KA)
B1 Knowles Center (KN)
B1 Lake Hall (LA)
B2 Latinx Student Cultural Center (LC)
C1 Marino Recreation Center (MC)
D3 Matthews Arena (MA)
B2 Meserve Hall (ME)
C2 Mugar Life Sciences Building (MU)
B2 Nightingale Hall (NI)
A3 Renaissance Park (RP)
C2 Richards Hall (RI)
C2 Robinson Hall (RB)
C2 ROTC Office (RO)
A2 Ryder Hall (RY)
B2 Shillman Hall (SH)
B2 Snell Engineering Center (SN)
B2 Snell Library (SL)
C3 Badger & Rosen SquashBusters (SB)
B1 Stearns Center (ST)
A1 Visitor Center (VC)
B1 West Village F, G, H (WV)

Residence Buildings

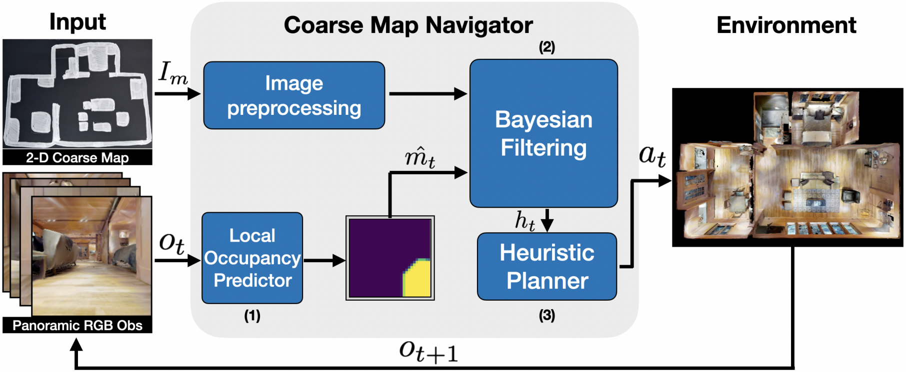
B3 10 Coventry (CV)
C1 106–122 Levine (106-122)
C1 142–148 Hemenway St (142)
C1 153 Hemenway St (LF)
C2 319 Huntington Ave (319)
C2 337 Huntington Ave (337)
B1 407 Huntington Ave (407)
B3 768 Columbus Ave (768)
B3 780 Columbus Ave (780)
B1 Burstein Hall (BU)
C3 Davenport Commons A, B (DC)
C2 East Village (EV)
C2 Hastings Hall (HS)
A2 International Village (IV)
D1 Kennedy Hall (KY)
C1 Kerr Hall (KH)
D1 Light Hall (LH)
C1 Loftman Hall (LF)
C1 Melvin Hall (MH)
B1 Rubenstein Hall (RU)
D1 Smith Hall (SM)
C1 Speare Hall (SP)
C1 Stetson East (SE)
C1 Stetson West (SW)
A1 West Village A, B, C, E (WV)
B1 West Village F, G, H (WV)
B1 White Hall (WH)
B1 Willis Hall (WI)

Parking Garages

C3 Columbus Parking Garage (CPG)
E2 Belvidere Parking Garage (BPG)
C2 Gainsborough Parking Garage (GPG)
A2 Renaissance Park Parking Garage (RPG)
A2 West Village Parking Garage (WPG)

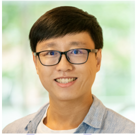
Northeastern University

Coarse map navigator

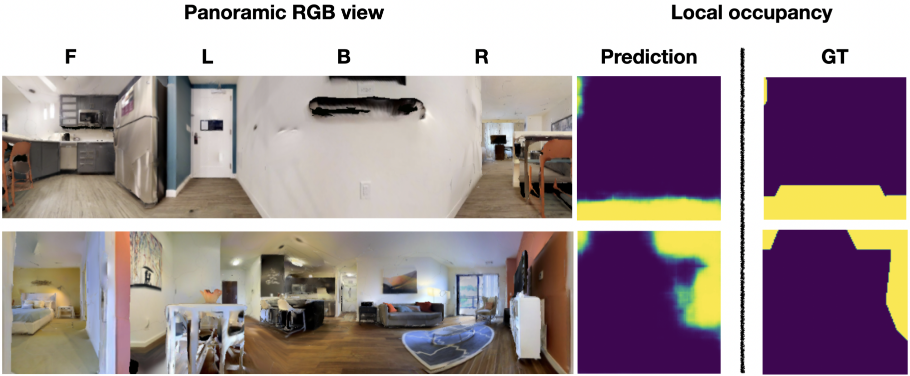


Chengguang Xu *et al.* Hierarchical robot navigation in novel environments using rough 2-D maps. CoRL 2020.

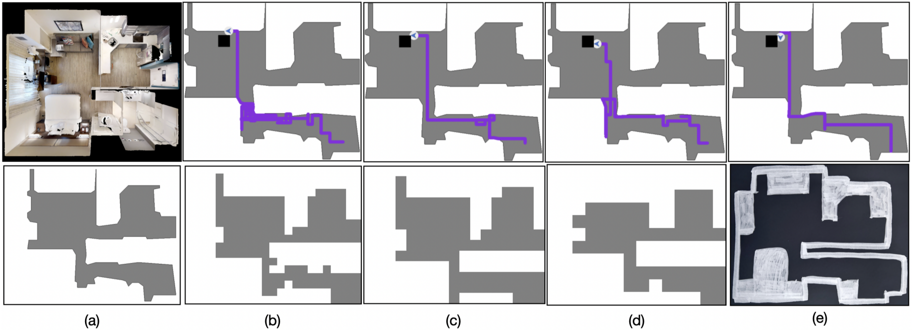
Chengguang Xu *et al.* Robot navigation in unseen environments using coarse maps. In submission.



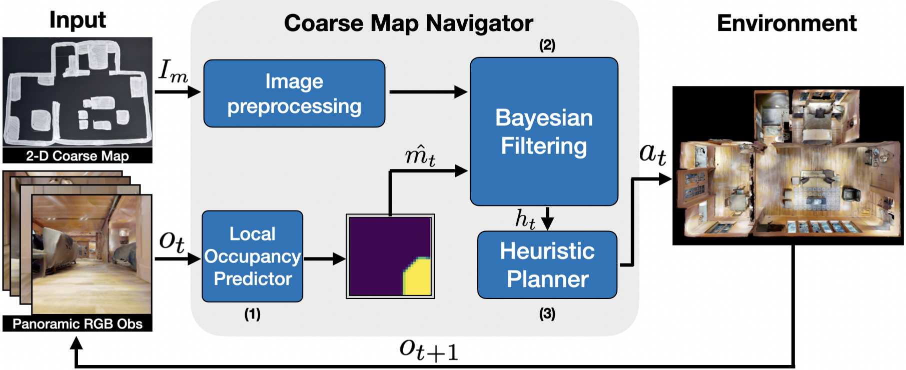
Coarse map navigator



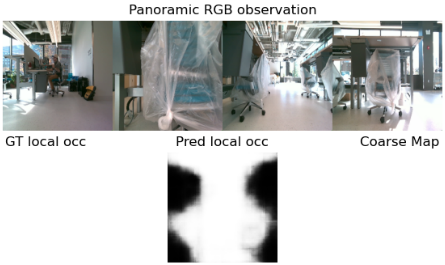
Coarse map navigator

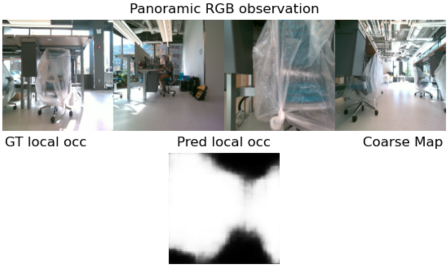


Opportunities for leveraging symmetry

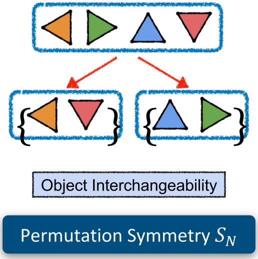


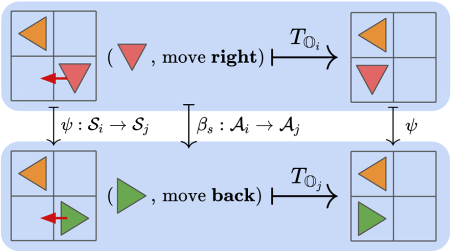
Opportunities for leveraging symmetry





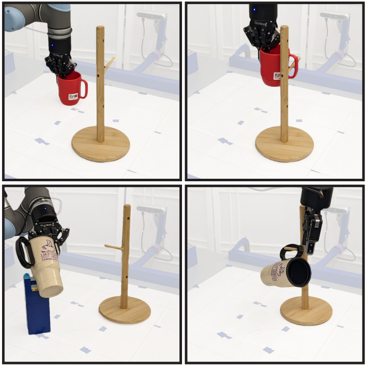
Why does CMN work? Compositional generalization





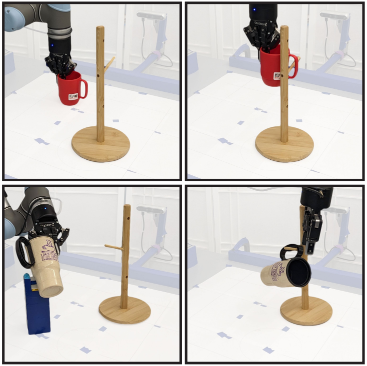
Linfeng Zhao *et al.* Toward compositional generalization in object-oriented world modeling. ICML 2022.

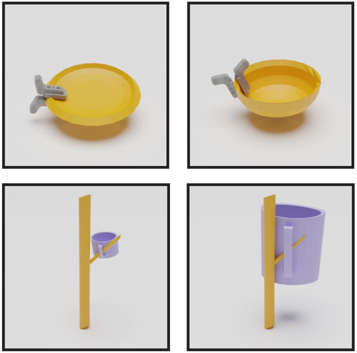
Following one high-level trajectory



Ondrej Biza *et al.* One-shot imitation learning via interaction warping. CoRL 2023.

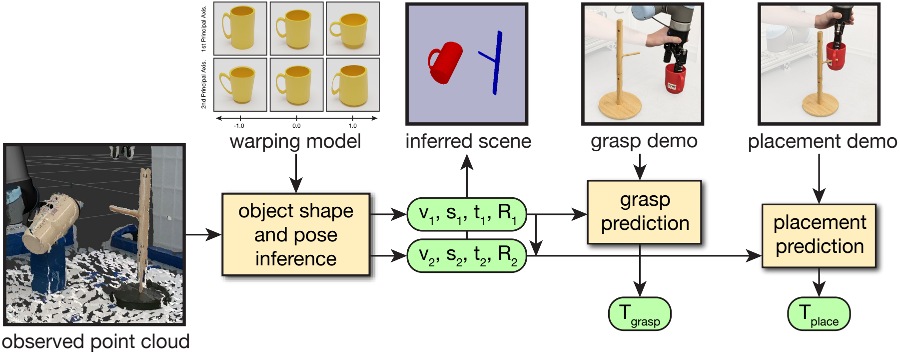
Following one high-level trajectory





Ondrej Biza *et al.* One-shot imitation learning via interaction warping. CoRL 2023.

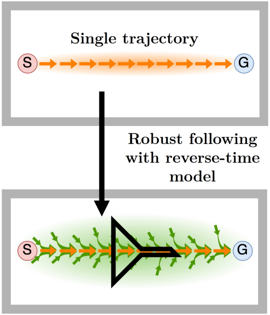
Following one high-level trajectory



Symmetry in RL and decision making

- Exploitation
  Integrating symmetry into differentiable planning
- Opportunity
  Robot navigation with coarse maps
- **Exploration**
  Speculative wish list

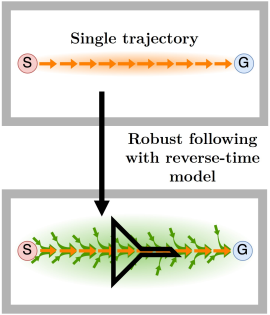
Demonstration × symmetry

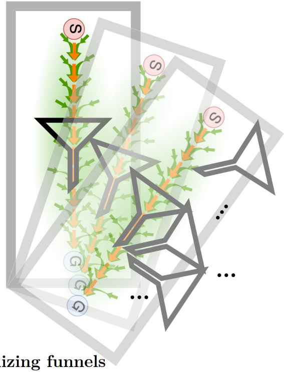


Jung Yeon Park & Wong. Robust imitation of a few demonstrations with a backwards model. NeurIPS 2022.

See John's poster! (on a different subject)

Demonstration × symmetry





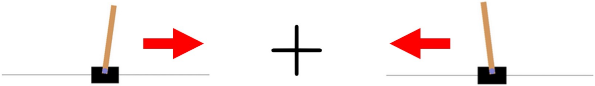
Generalizing funnels
with symmetry

Jung Yeon Park & Wong. Robust imitation of a few demonstrations with a backwards model. NeurIPS 2022.

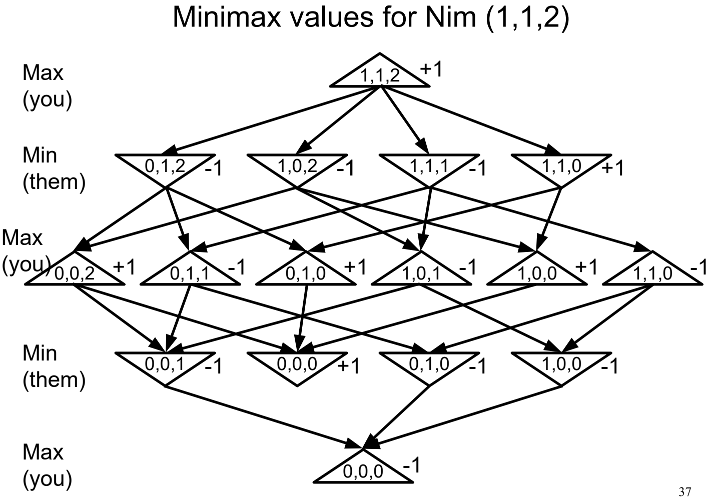
See John's poster! (on a different subject)

Symmetry discovery by demonstration

Discovering symmetry by demonstration:
Inferring reflection symmetry in cart-pole

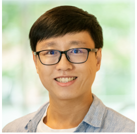


The symmetry of (repeated) computation



Acknowledgments and references

Linfeng Zhao



Chengguang Xu

Ondřej Bíža

Jung Yeon Park

Robin Walters

Christopher Amato

Robert Platt

Northeastern University
Khoury College of Computer Sciences

NSF National Science Foundation

Xupeng Zhu, Lingzhi Kong, Hongyu Li, Owen Howell, Dian Wang

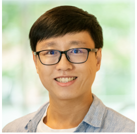
Kevin Robb, Skye Thompson, Kishore Reddy Pagidi, Abhinav Kumar, Neel Sortur

Taşkın Padır, Huaizu Jiang

Elise van der Pol, Thomas Kipf, Jan-Willem van de Meent

Acknowledgments and references


Linfeng Zhao


Chengguang Xu


Ondřej Bíža


Jung Yeon Park


Robin Walters


Christopher Amato


Robert Platt

Northeastern University
Khoury College of Computer Sciences

National Science Foundation

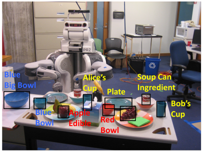
Integrating symmetry into differentiable planning with steerable convolutions. ICLR 2023

$E(2)$-equivariant graph planning for navigation. In submission.

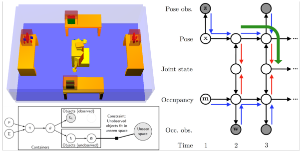
Can Euclidean symmetry be leveraged in RL and planning? In submission.

Hierarchical robot navigation in novel environments using rough 2-D maps. CoRL 2020.

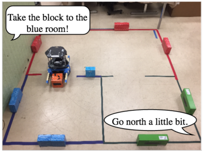
Robot navigation in unseen environments using coarse maps. In submission.



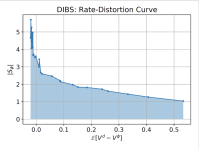
Object-based world modeling



Efficient state estimation

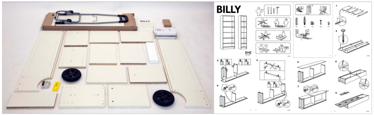


Grounding natural language



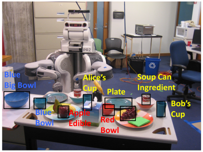
Fundamental limits of abstraction

Learning
Task-relevant
State representations

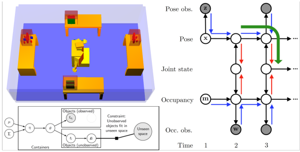


Learning to use existing abstractions

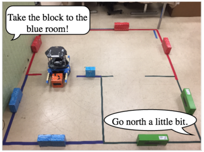
Let's chat! `lsw@ccs.neu.edu`



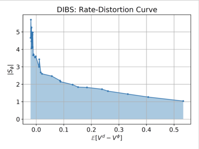
Object-based world modeling



Efficient state estimation

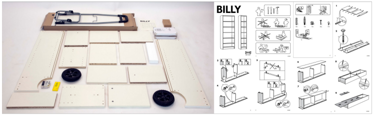


Grounding natural language



Fundamental limits of abstraction

Learning
Task-relevant
State representations



Learning to use existing abstractions

My research agenda

Identify and learn **intermediate state representations** that enable effective robot learning and planning, and therefore enable **robot generalization**.

Let's chat! `lsw@ccs.neu.edu`