Causality-Inspired Reinforcement Learning

State abstractions, exploration, and representations

Zizhao Wang

02/23/2026

Reinforcement Learning (RL)

RL method deficiencies

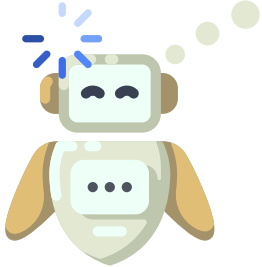
Most RL methods suffer from
- low sample efficiency
- poor generalization

Causes include three aspects:
- state space
- reward space
- action space

RL method deficiencies

Most RL methods suffer from
- low sample efficiency
- poor generalization

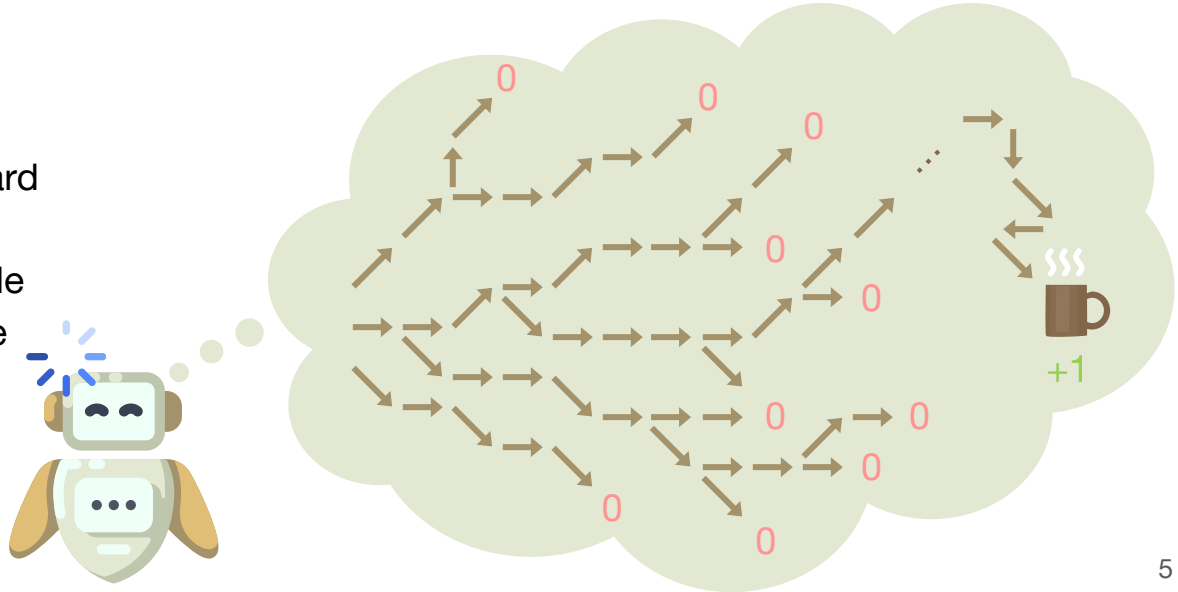
Causes include three aspects:
- state space
  - condition on all state factors
  - large input space → low sample efficiency
  - spurious correlations → poor generalization



RL method deficiencies

Most RL methods suffer from
- low sample efficiency
- poor generalization

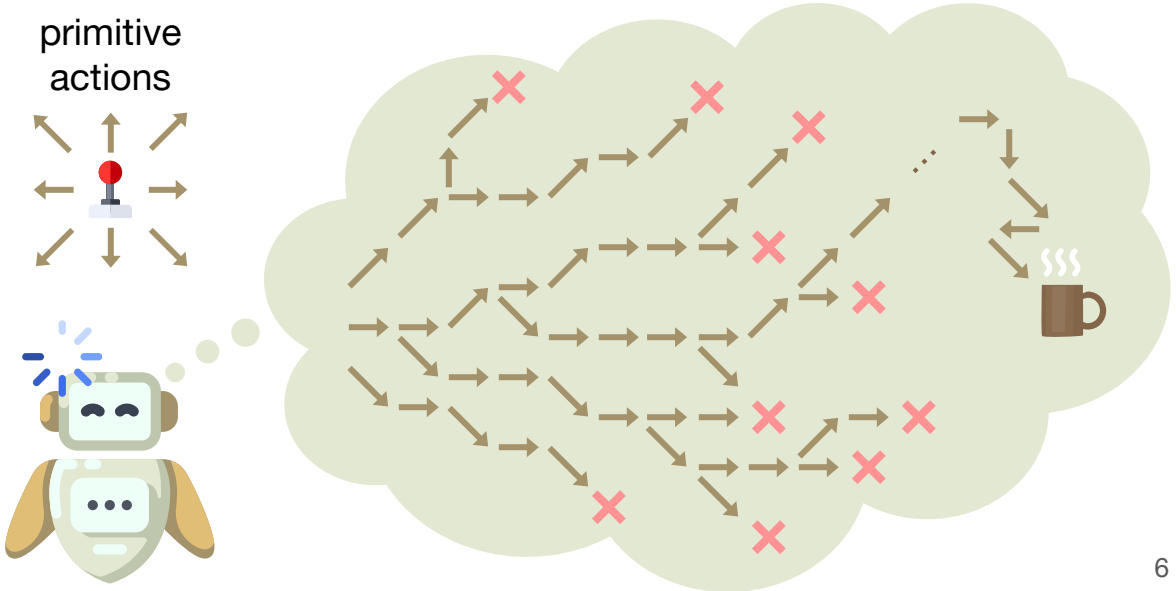
Causes include three aspects:
- state space
- reward space
  - Designing a dense reward is non-trivial.
  - Sparse rewards give little feedback → low sample efficiency.



RL method deficiencies

Most RL methods suffer from
- low sample efficiency
- poor generalization

Causes include three aspects:
- state space
- reward space
- action space
  - Primitive actions are inefficient.
  - Exploration space grows exponentially → low sample efficiency.



Causality-inspired RL

State Abstractions

focus on task-relevant state factors

ICML 2022, AAAI 2024

Intrinsic Rewards

provide extra exploration signals

NeurIPS 2023

Unsupervised Skill Discovery

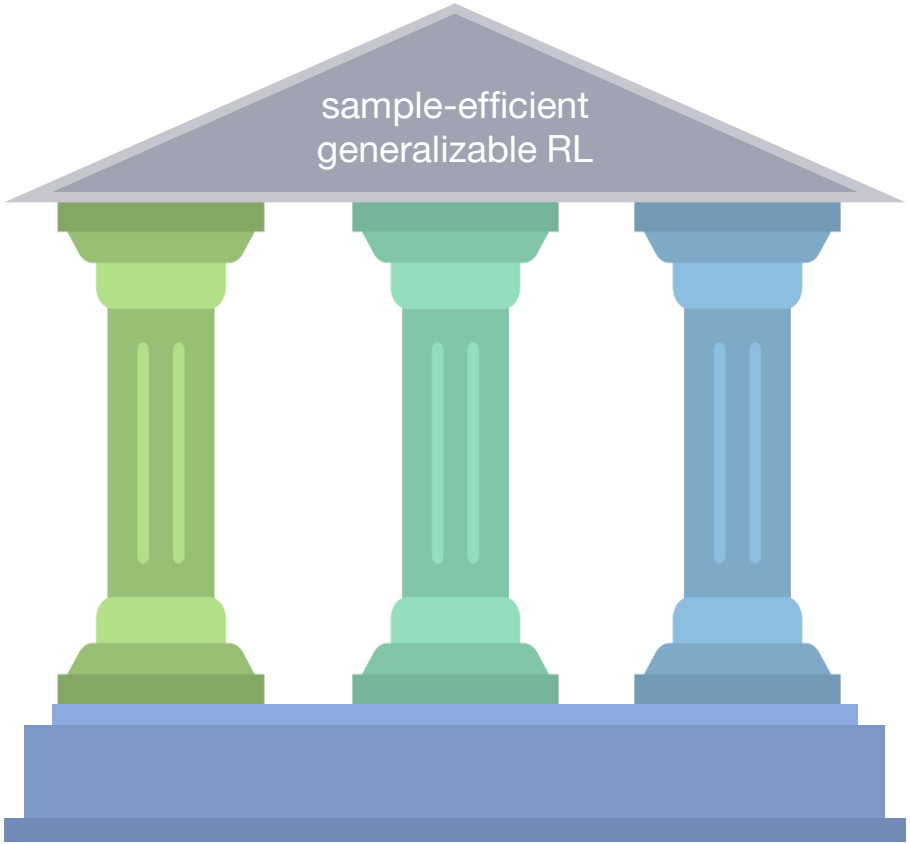
reduces exploration horizons

NeurIPS 2024

Representation Learning

extracts state factors from observations

NeurIPS 2025, preprint



Reinforcement Learning (RL)

factored Markov decision processes

$$M = (S, O, A, R, P, \gamma)$$

S: factored state space, consisting of K state factors

$$S = S^1 \times \cdots \times S^K$$

O: observation space

A: action space

R: reward function

P: transition probability (dynamics)

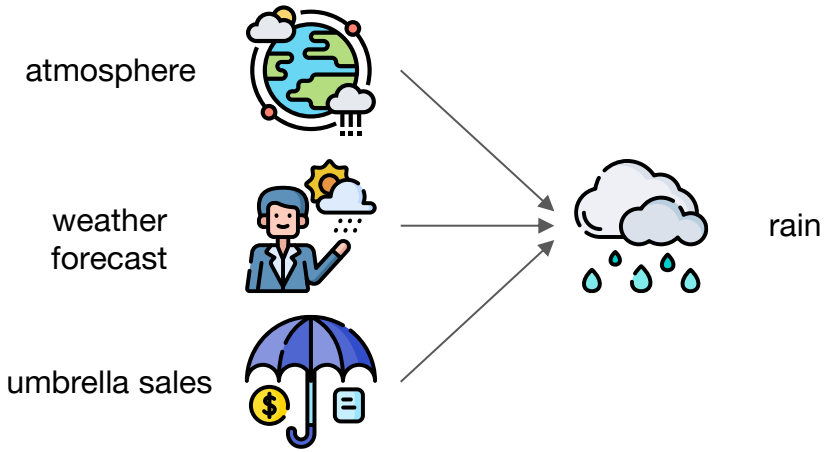
γ: discount factor

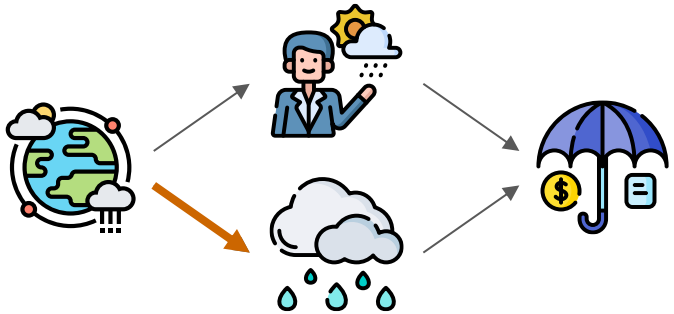
Causality

End-to-end learning:
- uses all inputs to make predictions, no matter if an input is necessary.

Granger causality:
- uses an input if it increases the prediction performance.

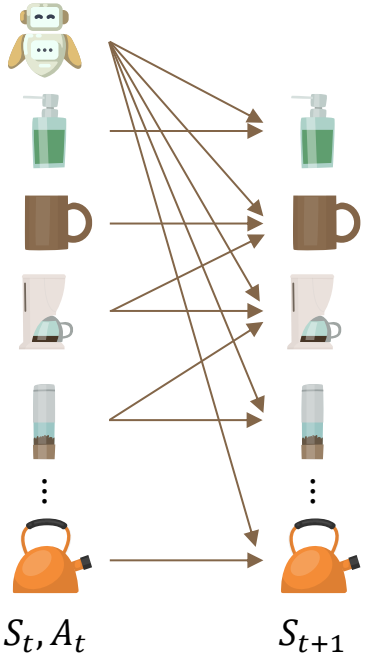
Causality:
- Even if an input helps prediction, it is not necessarily a direct cause (could be a child / sibling / grandparent …).
- Causality identifies true causes with additional assumptions / constraints.
- Conditioning only on true causes facilitates generalization.

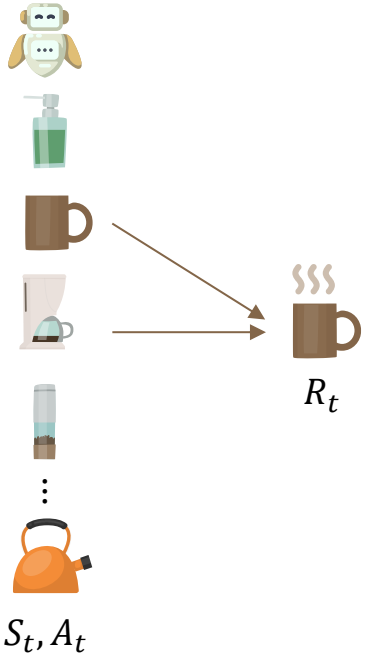




Causality in factored MDPs

Our work mainly focuses on causal relationships in the dynamics and the reward function.





Causality-inspired RL

State Abstractions

focus on task-relevant state factors

ICML 2022, AAAI 2024

Intrinsic Rewards

provide extra exploration signals

NeurIPS 2023

Unsupervised Skill Discovery

reduces exploration horizons

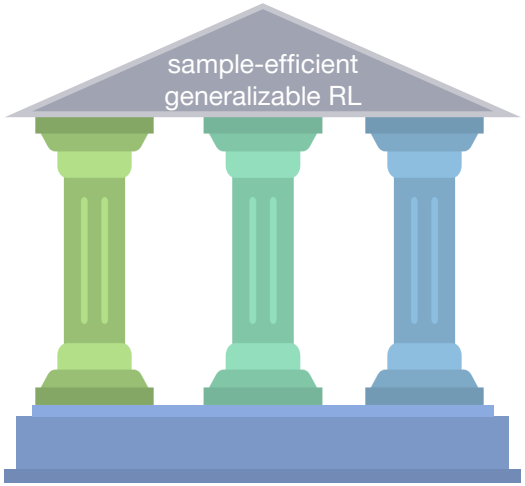
NeurIPS 2024

assume state factors are given

Representation Learning

extracts state factors from observations

NeurIPS 2025, preprint



Causality-inspired RL

State Abstractions

focus on task-relevant state factors

ICML 2022, AAAI 2024

Intrinsic Rewards

provide extra exploration signals

NeurIPS 2023

Unsupervised Skill Discovery

reduces exploration horizons

NeurIPS 2024

Representation Learning

extracts state factors from observations

NeurIPS 2025, preprint

Causality-inspired RL

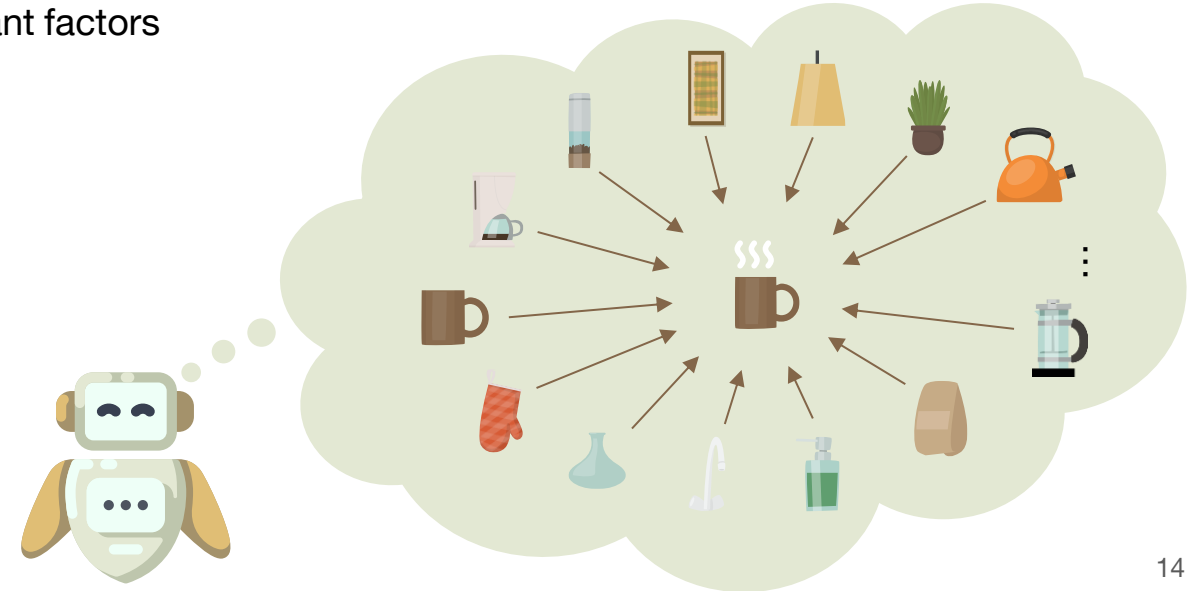
State space

- condition on all state factors
- large input space
- spurious correlations

Causality-inspired RL

State space
- condition on all state factors

State abstractions
- condition only on task-relevant factors

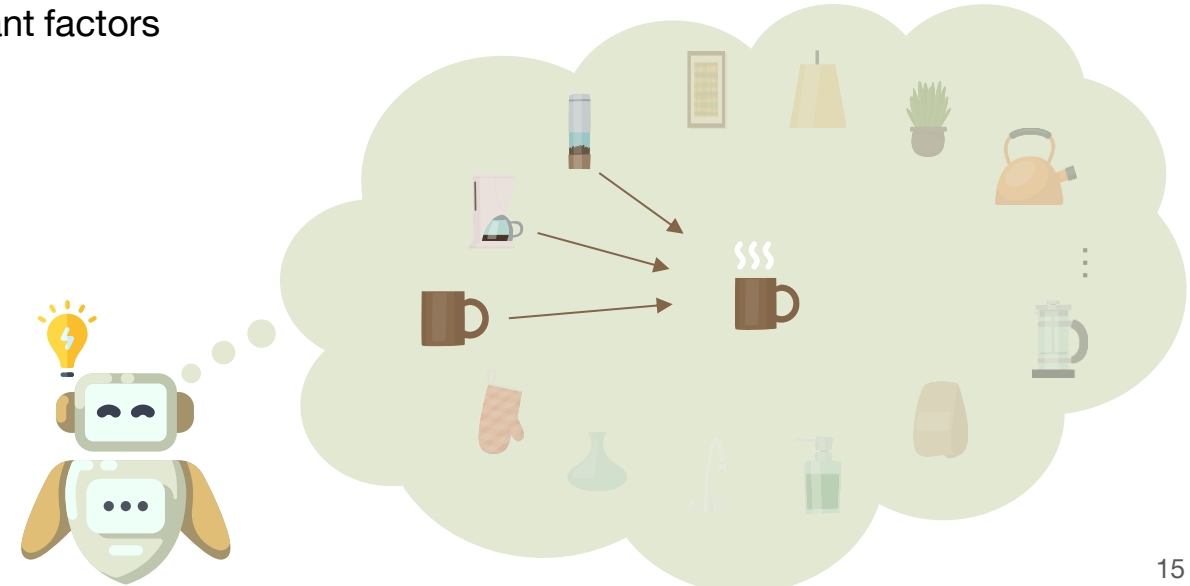


Wang et al., CDL, ICML 2022. Wang et al., CBM, AAAI 2024

Causality-inspired RL

State space
- condition on all state factors

State abstractions
- condition only on task-relevant factors
- reduce learning space
- facilitate generalization



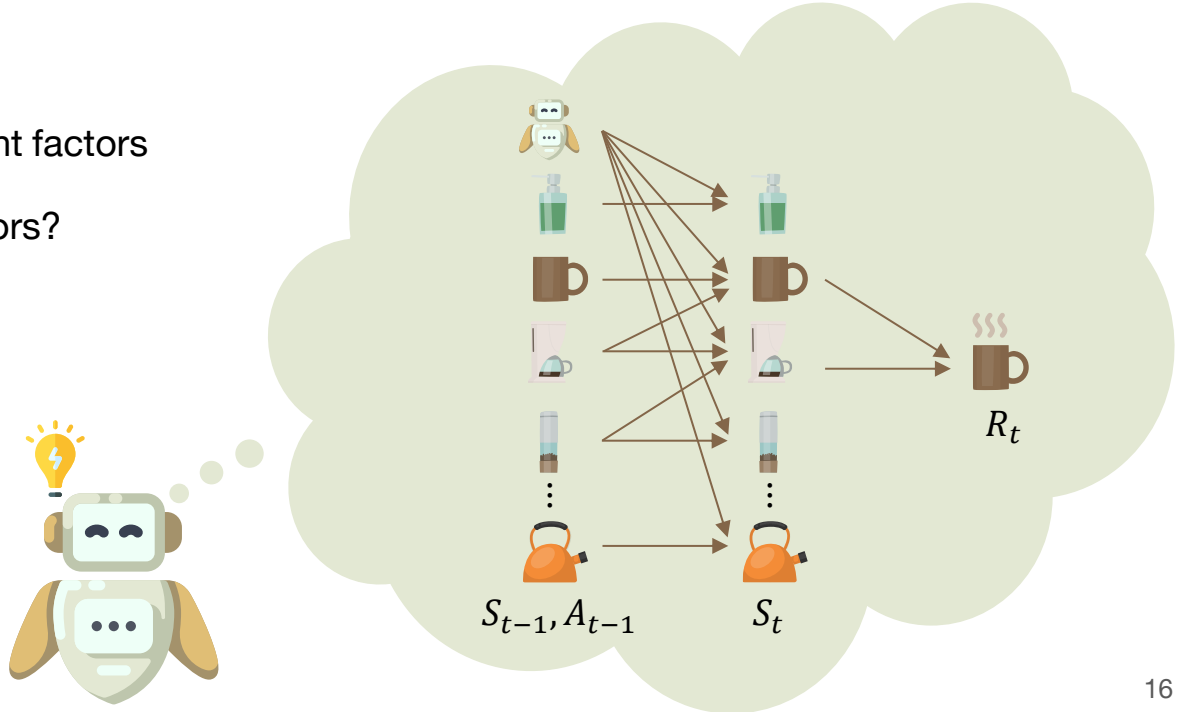
Wang et al., CDL, ICML 2022. Wang et al., CBM, AAAI 2024

Causality-inspired RL

State space
- condition on all state factors

State abstractions
- condition only on task-relevant factors

How to identify task-relevant factors?



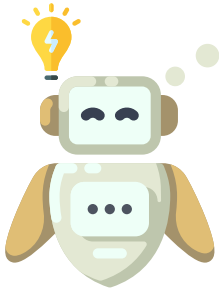
Wang et al., CDL, ICML 2022. Wang et al., CBM, AAAI 2024

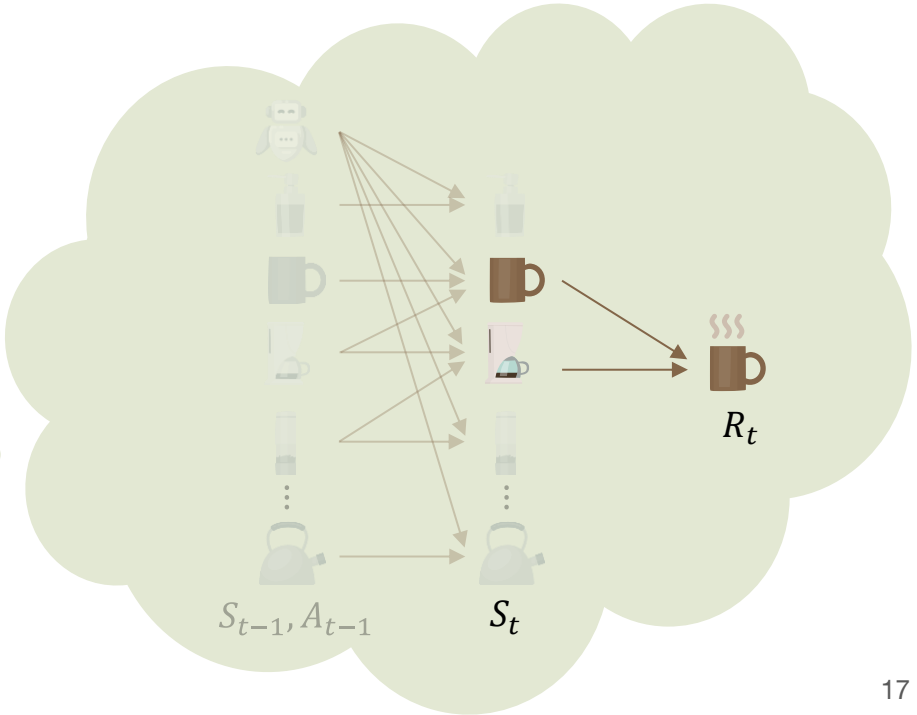
Causality-inspired RL

State space
- condition on all state factors

State abstractions
- condition only on task-relevant factors

How to identify task-relevant factors?
- identify reward function's causal parents





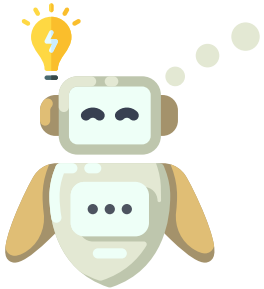
Wang et al., CDL, ICML 2022. Wang et al., CBM, AAAI 2024

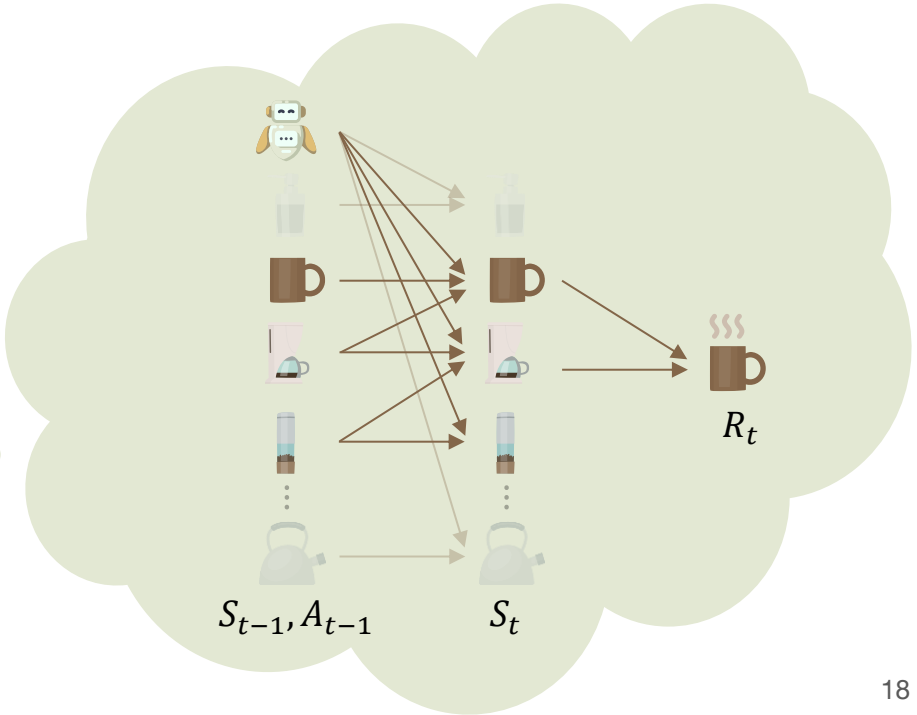
Causality-inspired RL

State space
- condition on all state factors

State abstractions
- condition only on task-relevant factors

How to identify task-relevant factors?
- identify reward function's causal parents
- identify their dynamics causal ancestors





Wang et al., CDL, ICML 2022. Wang et al., CBM, AAAI 2024

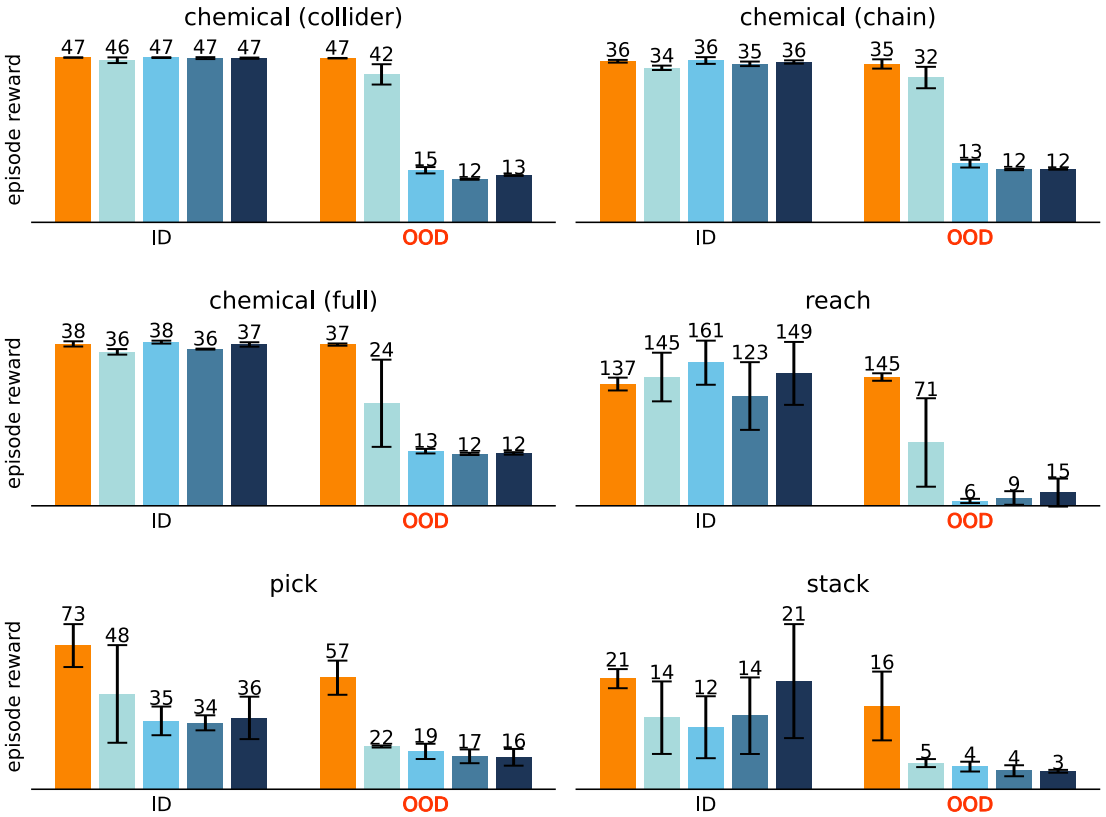
Causality-inspired RL

Causal state abstractions facilitates sample efficiency and generalization.

- Causal Dynamics Learning (Ours)
- Regularization
- Graph Neural Network
- Modular
- Monolithic

ID: in-distribution states

OOD: out-of-distribution states



Wang et al., CDL, ICML 2022. Wang et al., CBM, AAAI 2024

Causality-inspired RL

State Abstractions

focus on task-relevant state factors

ICML 2022, AAAI 2024

Intrinsic Rewards

provide extra exploration signals

NeurIPS 2023

Unsupervised Skill Discovery

reduces exploration horizons

NeurIPS 2024

Representation Learning

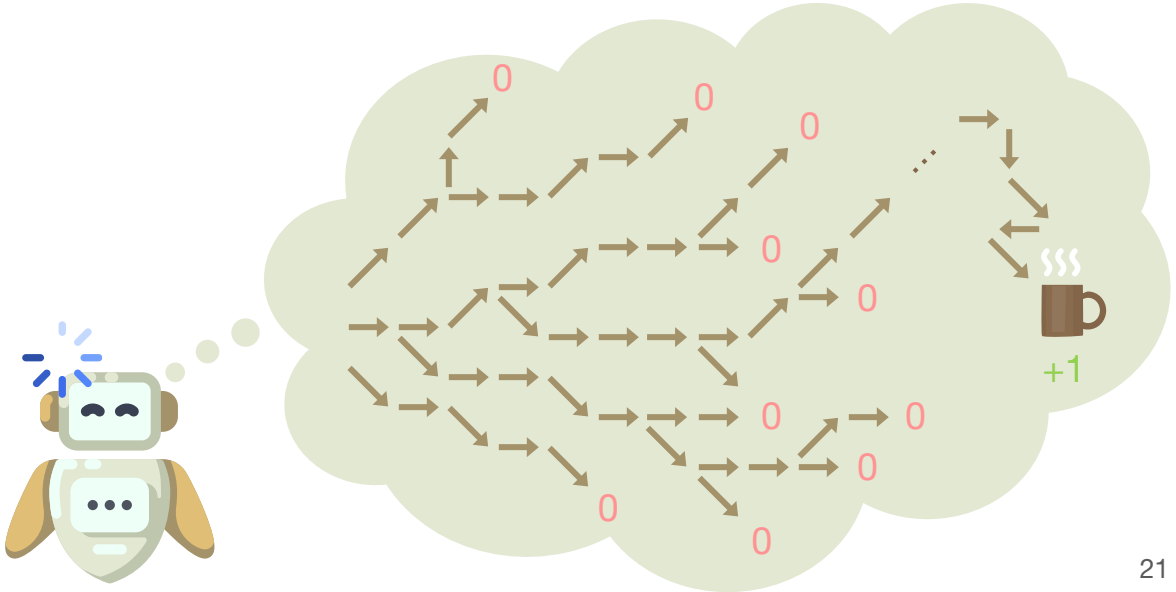
extracts state factors from observations

NeurIPS 2025, preprint

Causality-inspired RL

Reward space

- Designing a dense reward is non-trivial.
- Sparse rewards give little feedback.



Wang et al., ELDEN, NeurIPS 2023.

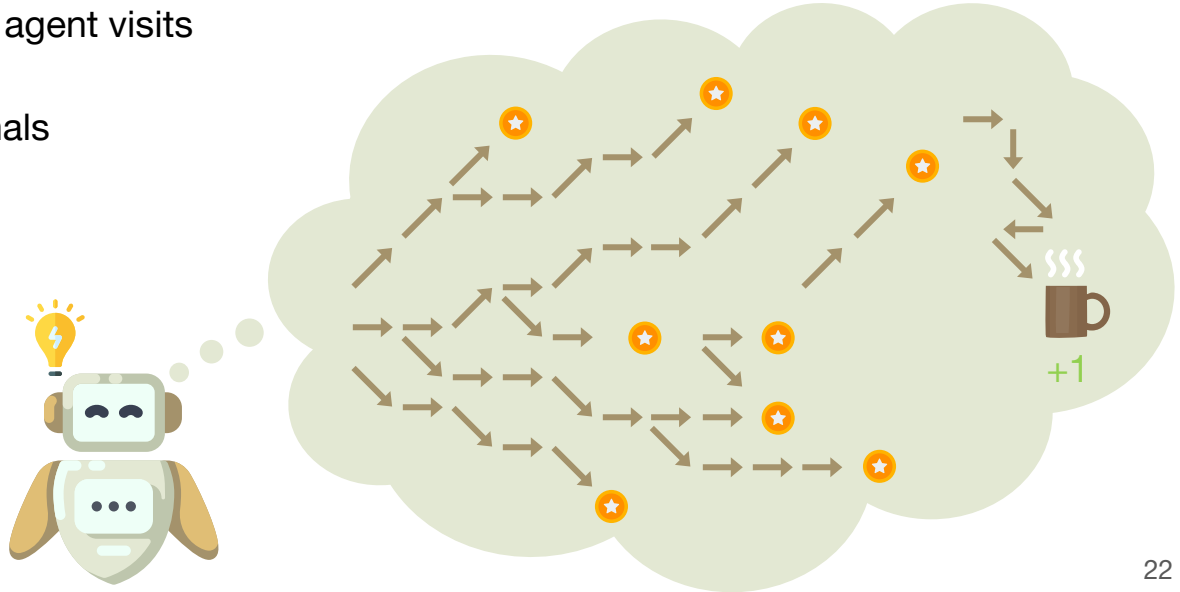
Causality-inspired RL

Reward space

- Sparse rewards give little feedback.

Intrinsic rewards

- extra reward bonus when the agent visits “interesting” states
- provide extra exploration signals



Wang et al., ELDEN, NeurIPS 2023.

Causality-inspired RL

Reward space

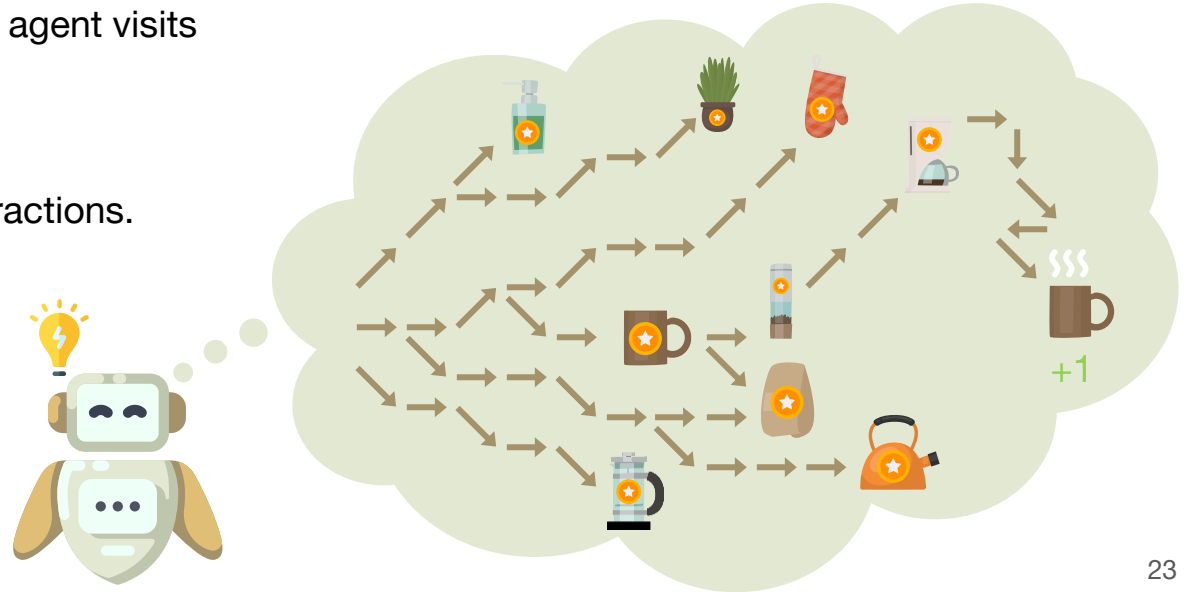
- Sparse rewards give little feedback.

Intrinsic rewards

- extra reward bonus when the agent visits “interesting” states

What states are “interesting”?

- The agent induces novel interactions.



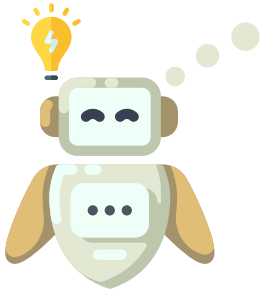
Wang et al., ELDEN, NeurIPS 2023.

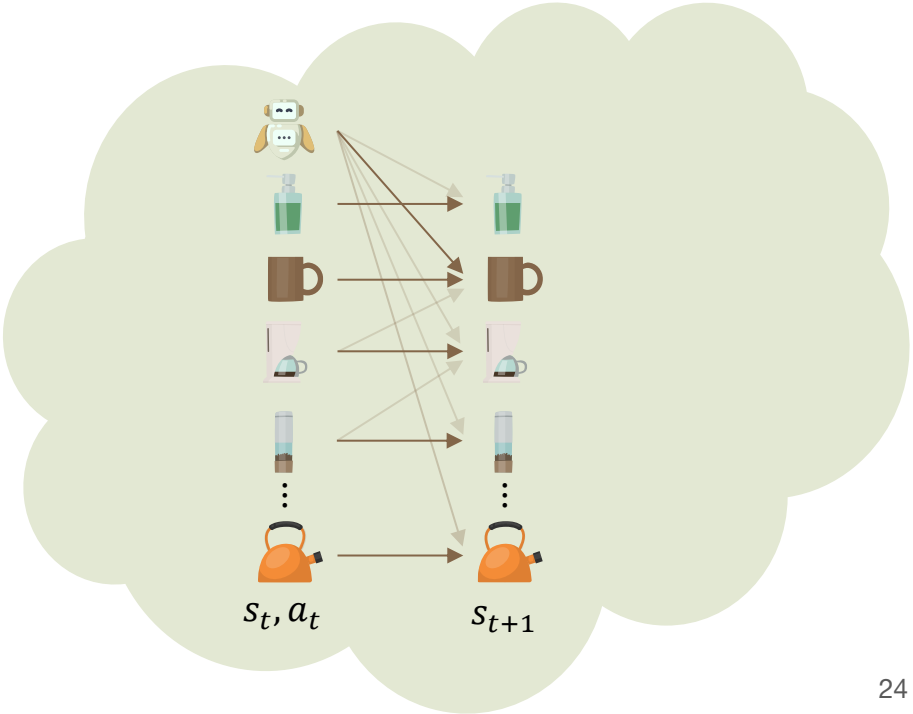
Causality-inspired RL

Reward space
- Sparse rewards give little feedback.

Intrinsic rewards
- extra reward bonus when the agent visits "interesting" states

What states are "interesting"?
- The agent induces novel interactions.
- Interactions are defined as **state-specific** causal dependencies.

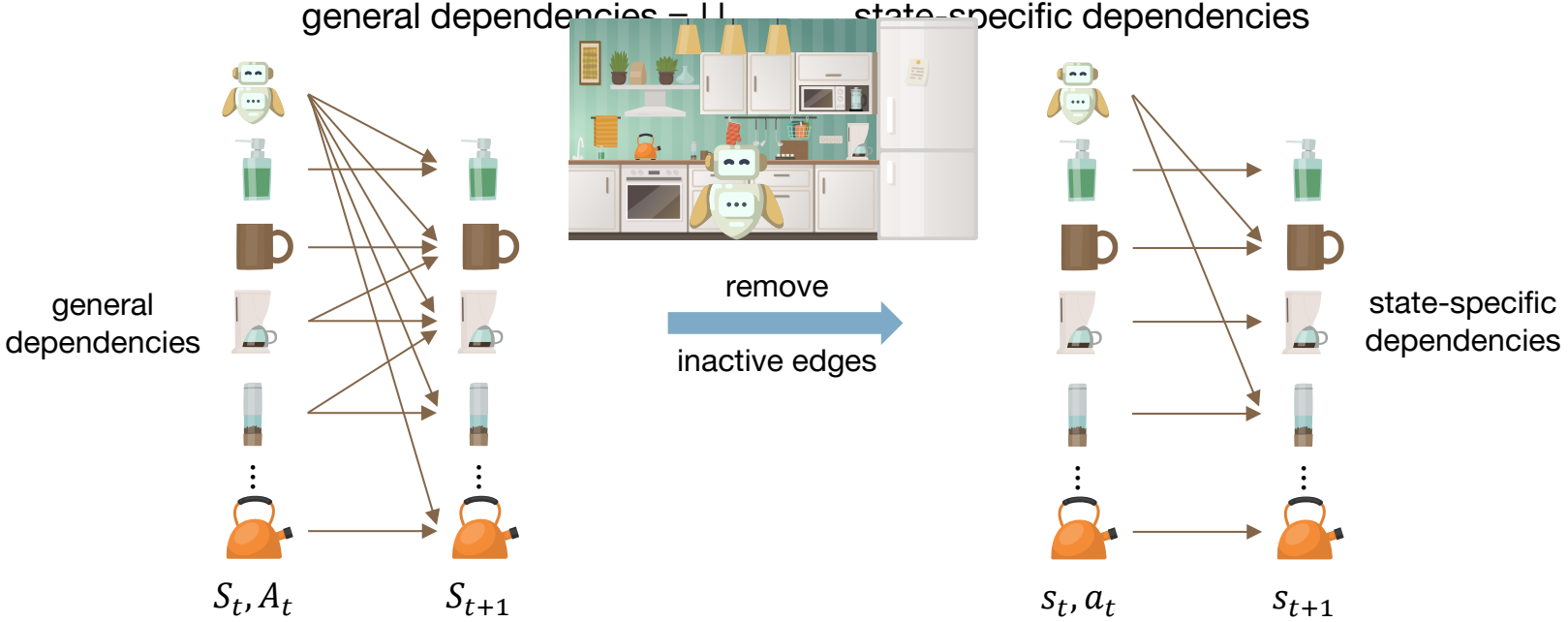




Wang et al., ELDEN, NeurIPS 2023.

General vs state-specific causal dependencies

General causal dependencies describe which influences are possible (A can influence B).
At a particular state, some influences are inactive.
State-specific dependencies are obtained by removing those inactive edges.



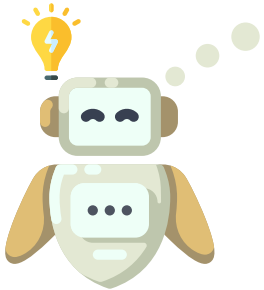
Wang et al., ELDEN, NeurIPS 2023.

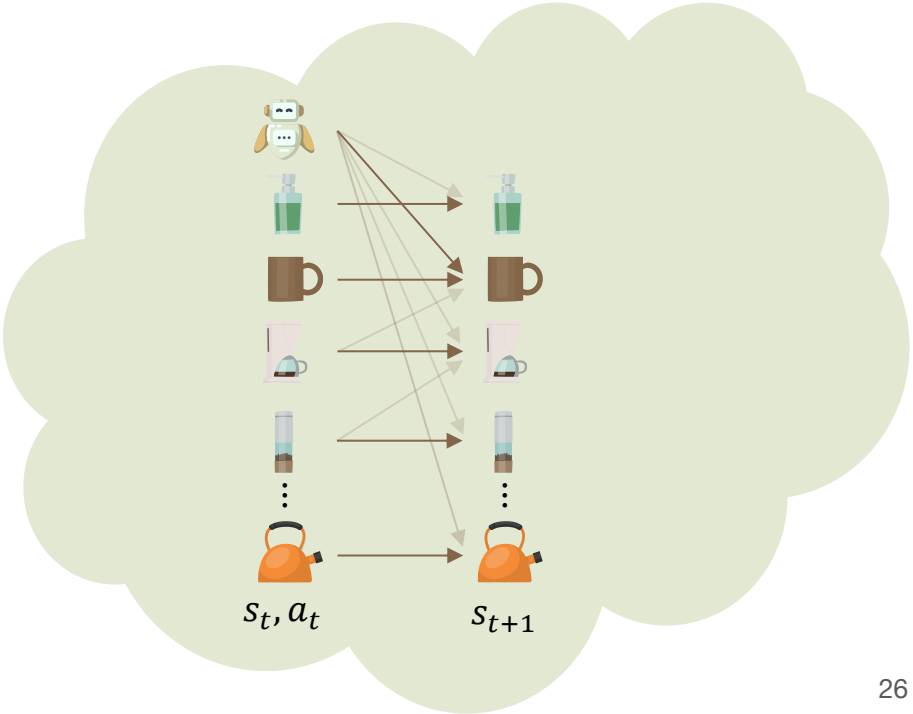
Causality-inspired RL

Reward space
- Sparse rewards give little feedback.

Intrinsic rewards
- extra reward bonus when the agent visits "interesting" states

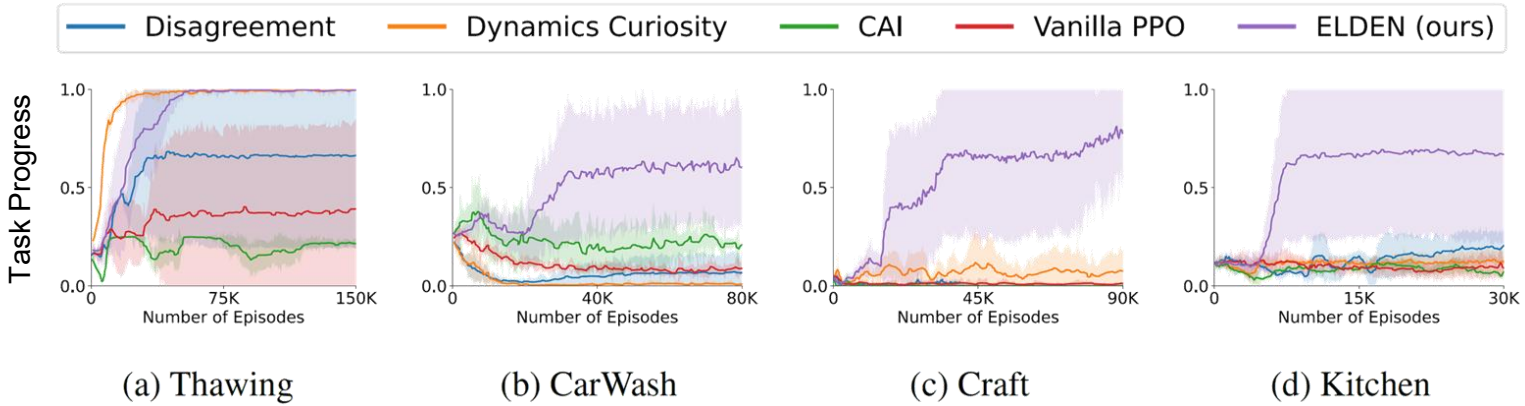
What states are "interesting"?
- The agent induces novel interactions.
- Interactions are defined as **state-specific** causal dependencies.
- Novelty is measured as uncertainty.





Wang et al., ELDEN, NeurIPS 2023.

Causality-inspired RL



ELDEN is more sample-efficient than learning without intrinsic rewards and than prior intrinsic reward methods.

Wang et al., ELDEN, NeurIPS 2023.

Causality-inspired RL

State Abstractions

focus on task-relevant state factors

ICML 2022, AAAI 2024

Intrinsic Rewards

provide extra exploration signals

NeurIPS 2023

Unsupervised Skill Discovery

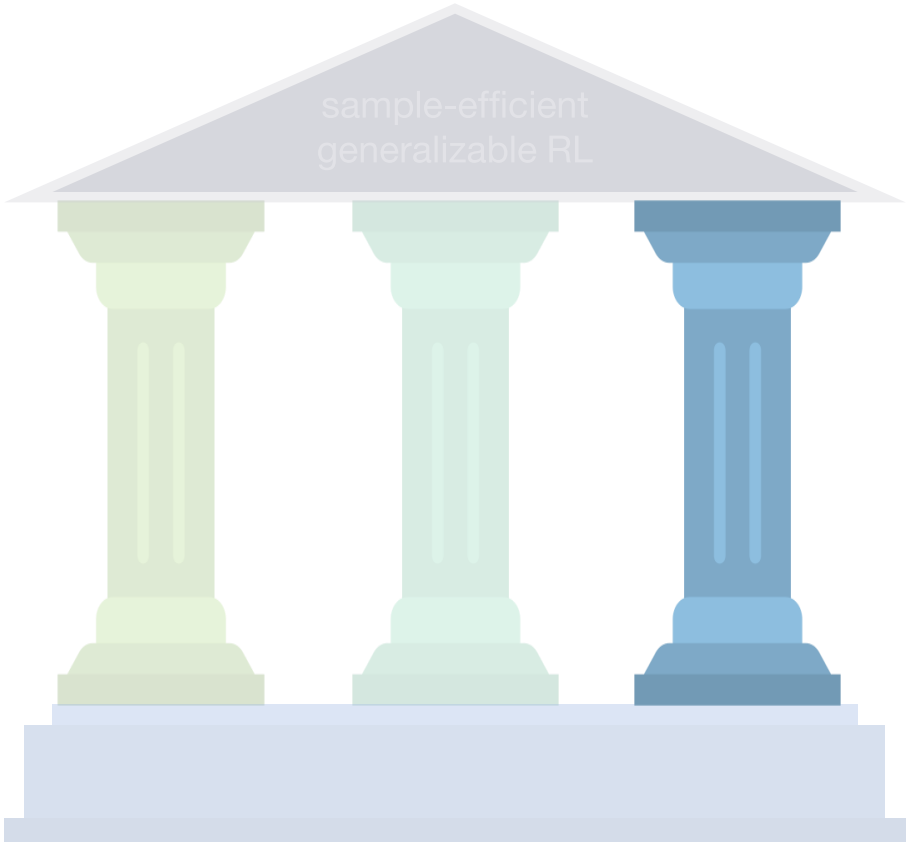
reduces exploration horizons

NeurIPS 2024

Representation Learning

extracts state factors from observations

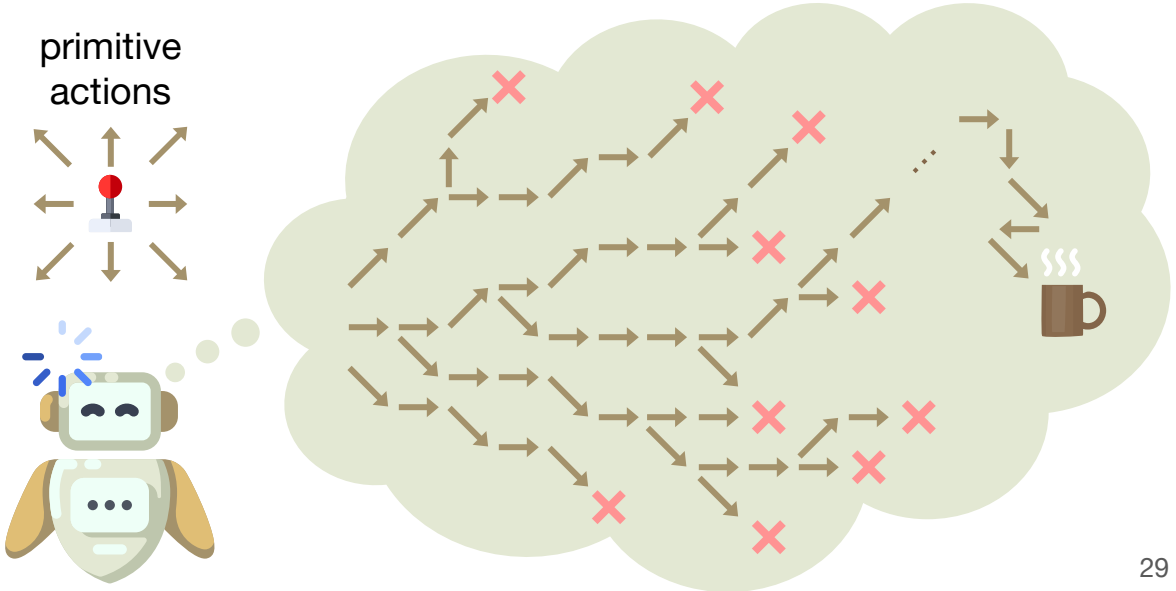
NeurIPS 2025, preprint



RL method deficiencies

Action space

- primitive actions are inefficient.
- exploration space grows exponentially with task horizons.



Wang et al., SkiLD, NeurIPS 2024.

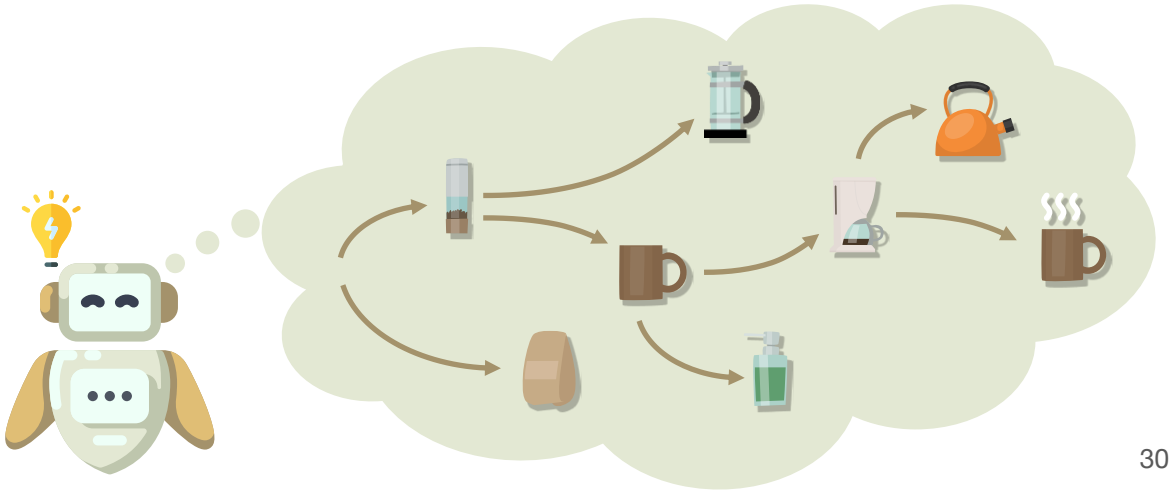
Causality-inspired RL

Action space

- exploration space grows exponentially with task horizons.

Unsupervised skill discovery

- learns reusable skills for downstream tasks
- reduces exploration horizons



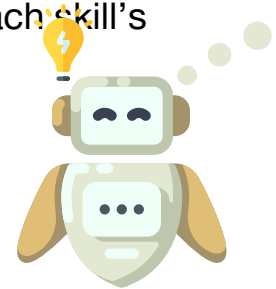
Wang et al., SkiLD, NeurIPS 2024.

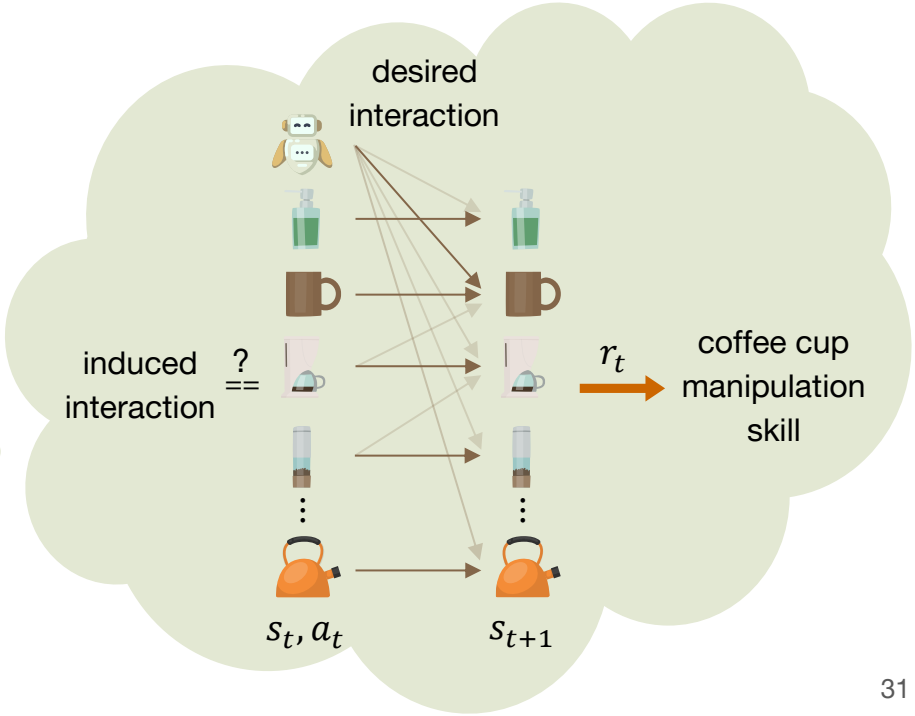
Causality-inspired RL

Action space
- exploration space grows exponentially with task horizons.

Unsupervised skill discovery
- learns reusable skills for downstream tasks

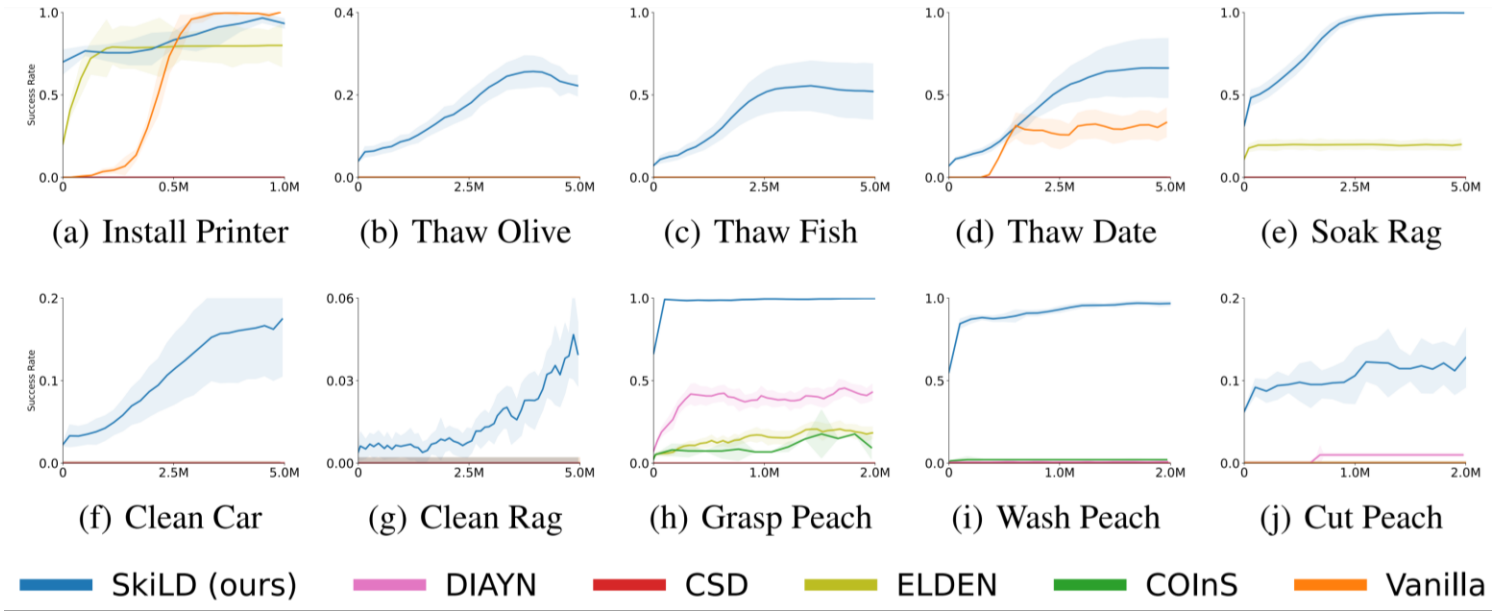
What skills to learn?
- the skills that induce diverse interactions
- use induced state-specific causal dependencies to compute each skill's rewards





Wang et al., SkiLD, NeurIPS 2024.

Causality-inspired RL



SkiLD is more sample-efficient than learning without skills and than prior skill discovery or intrinsic reward methods.

Wang et al., SkiLD, NeurIPS 2024.

Causality-inspired RL

State Abstractions

focus on task-relevant state factors

ICML 2022, AAAI 2024

Intrinsic Rewards

provide extra exploration signals

NeurIPS 2023

Unsupervised Skill Discovery

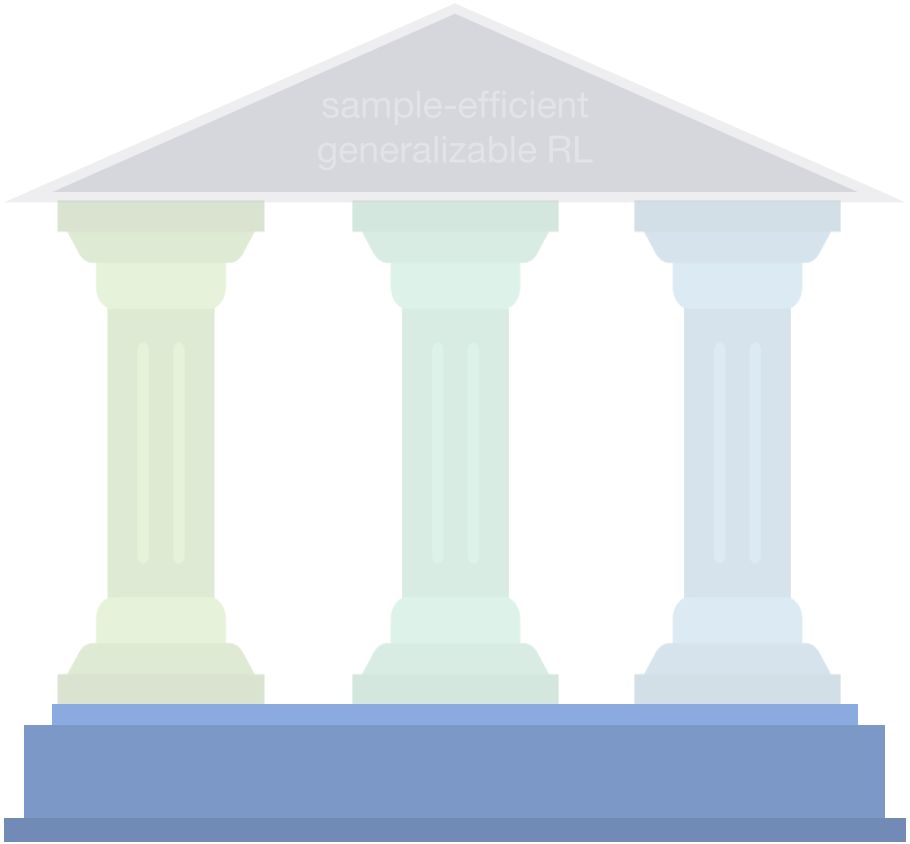
reduces exploration horizons

NeurIPS 2024

Representation Learning

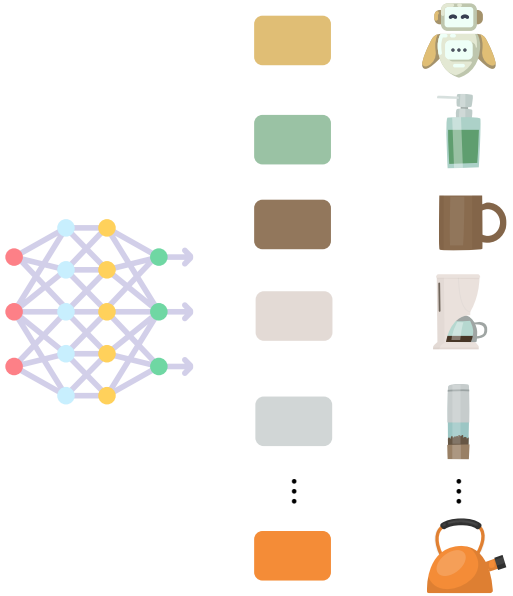
extracts state factors from observations

NeurIPS 2025, preprint



Factored representation learning

If we only have access to low-level observations, such as images, how can we extract state factors from them?



factored representations

Background - latent action models (LAMs)

In many real-world scenarios, multiple state factors are controlled by (relatively) independent actions.

Can independent actions serve as learning signals for discovering factored representations?

What if action labels are not available? We can build on latent action models.

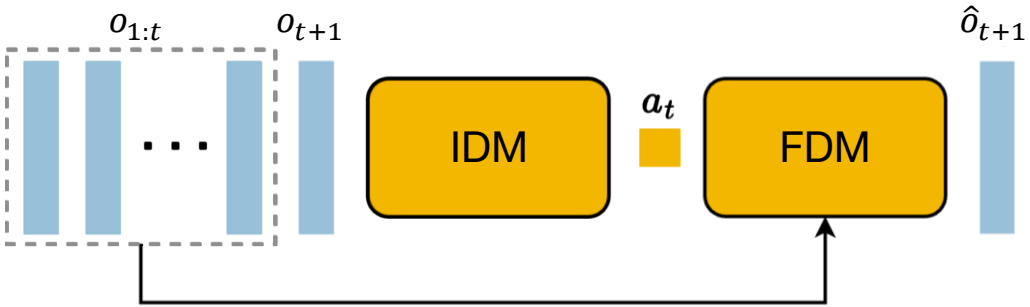
Background - latent action models (LAMs)

Latent action models aim to learn a world model without action labels.

Instead of predicting solely from observations $o_{1:t}$,

- an inverse dynamics model infers a latent action, $a_t = \mathrm{IDM}(o_{1:t}, o_{t+1})$;
- a forward dynamics model makes predictions, $\hat{o}_{t+1} = \mathrm{FDM}(o_{1:t}, a_t)$.

The model is trained to minimize the prediction error.

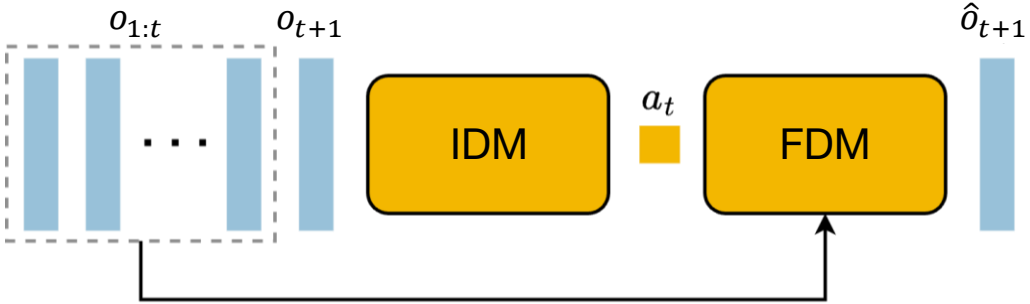


Bruce, Jake, et al. "Genie: Generative interactive environments." *ICML*. 2024.
Schmidt, Dominik, and Minqi Jiang. "Learning to act without actions." *ICLR*. 2024.

Background - latent action models (LAMs)

Why are LAMs about actions, rather than just ‘hindsight video models’? A key design choice is to prevent a shortcut:

- If a_t has unlimited capacity, it could simply copy o_{t+1} and bypass dynamics learning.
- To avoid that shortcut, a_t is usually constrained with an information bottleneck.
  - discrete codebooks
  - VAE-style regularization
- a_t should capture the info that’s unpredictable from $o_{1:t}$ (which usually is the agent’s action).



Bruce, Jake, et al. "Genie: Generative interactive environments." *ICML*. 2024.
Schmidt, Dominik, and Minqi Jiang. "Learning to act without actions." *ICLR*. 2024.

Motivation

This information bottleneck on latent actions could also bring a potential problem…

Do you see an issue in Genie 3's demo?

None of the people are moving,
unless they are prompted to do so.

Reason: the latent action has limited capacity to model complex joint movements.

https://deepmind.google/discover/blog/genie-3-a-new-frontier-for-world-models/

Motivation

However, in many scenarios, multiple state factors have their own actions.

- It's challenging to squeeze all action combinations into a single constrained action space.

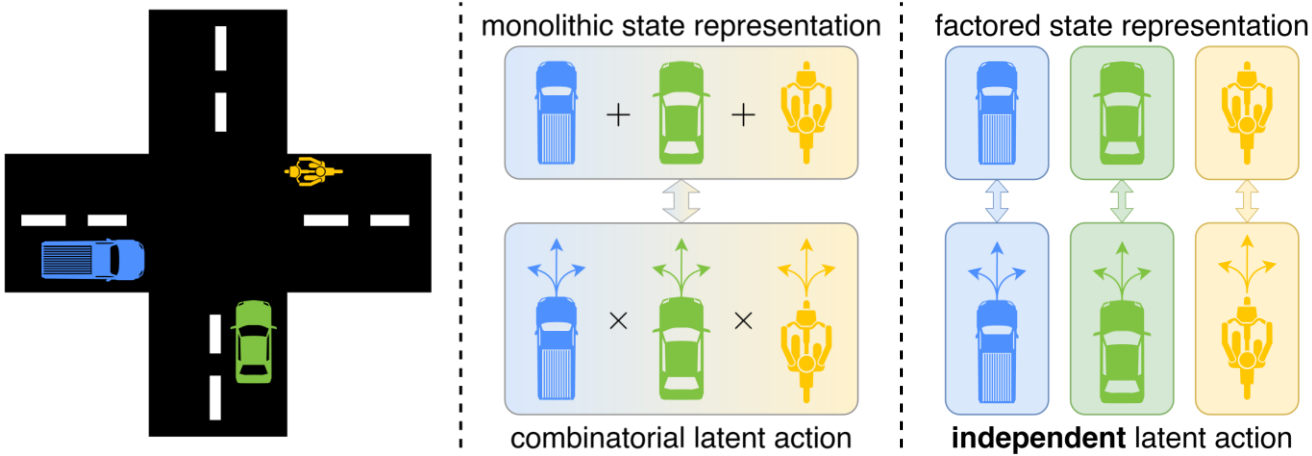
How can we handle this exponentially increasing action space?

- Factorization, which also provides a way to learn factored representations.

Method

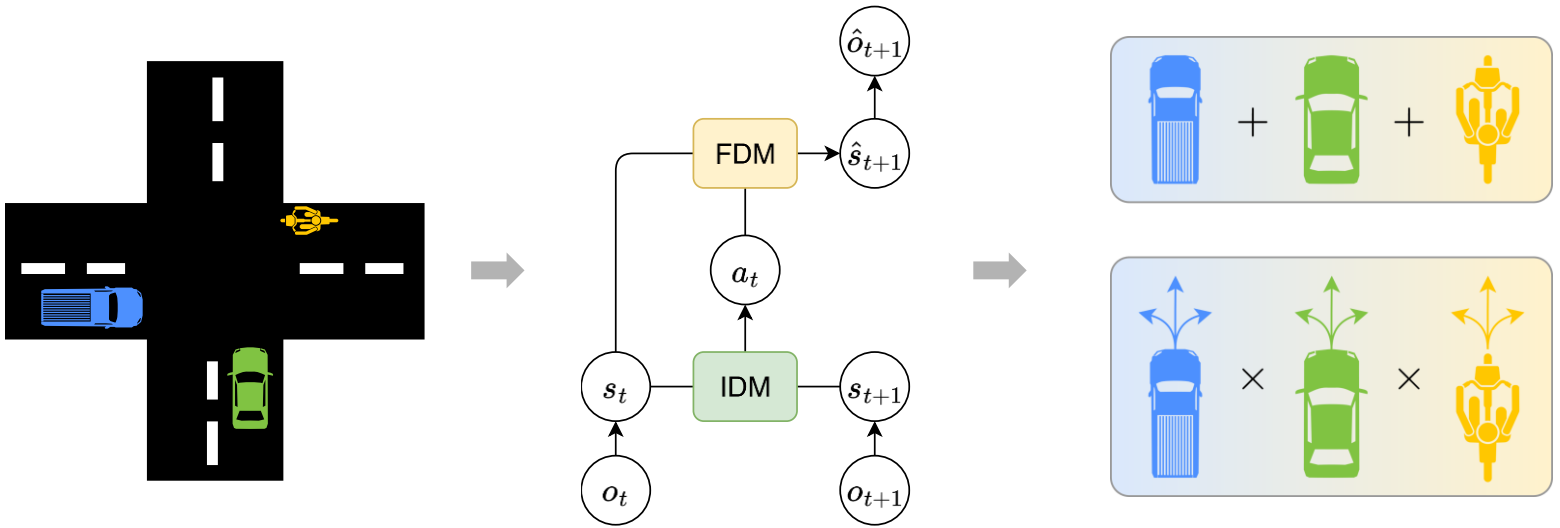
state & action representations: monolithic → factored

- The state representation consists of K factors.
- Each state factor has its own latent action.



Method

state & action representations: monolithic → factored

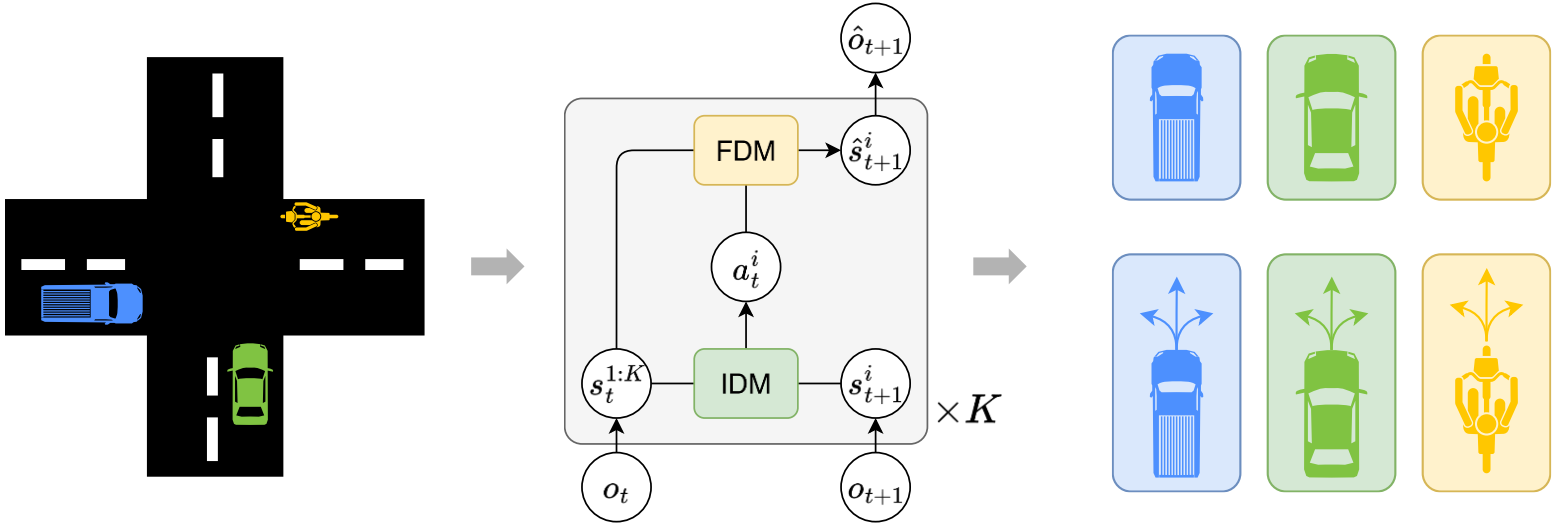


observation

monolithic LAM

Method

state & action representations: monolithic → factored



observation

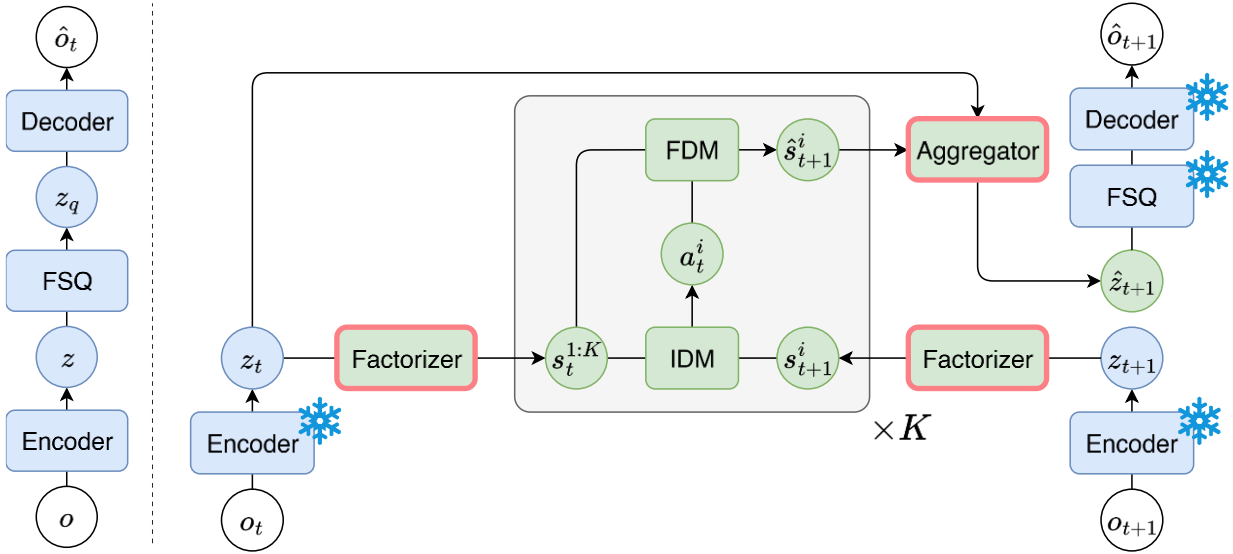
factored LAM

shared action space

Method

Training consists of two stages:

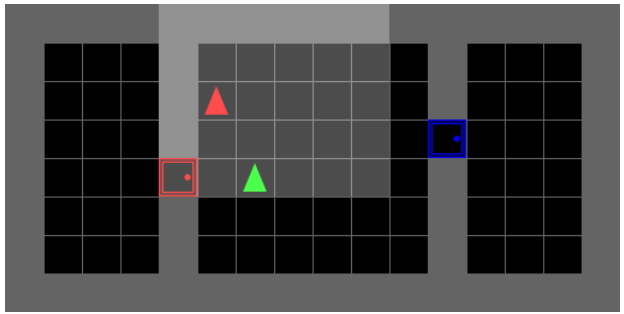
- Pre-train an encoder with the reconstruction loss, so we can learn the LAM in the feature space.
- Train the factored LAM with the prediction loss $\|z_{t+1} - \hat{z}_{t+1}\|$.
  - factorizer: split encoder features into factors, slot attention
  - aggregator: map factors back to features, cross-attn(query = z_t, key = $\hat{s}_{t+1}^{1:K}$)



Results

Datasets

- MultiGrid
- Procgen
- nuPlan

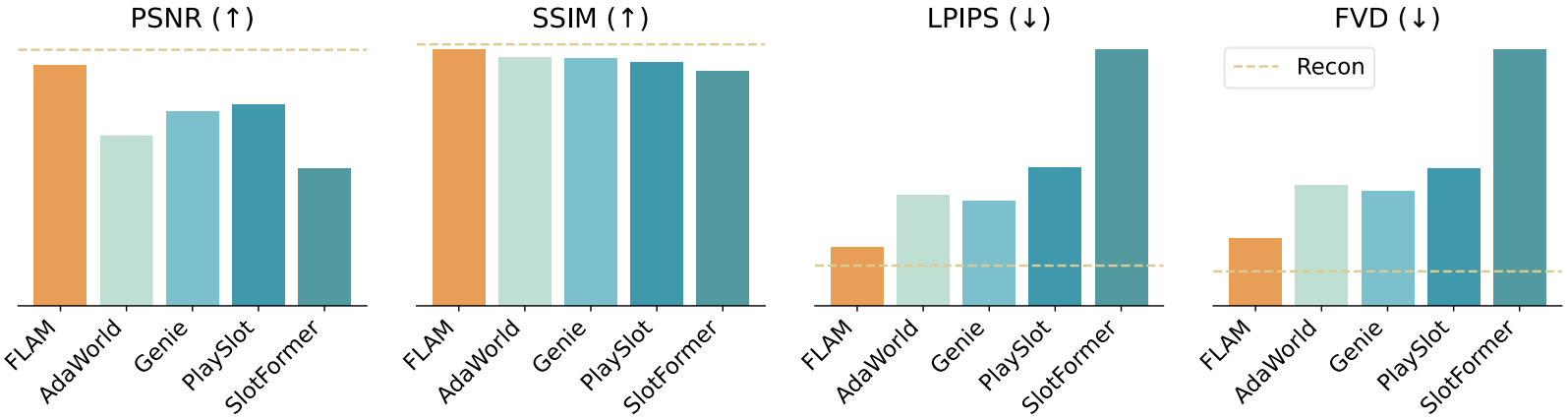


Results - prediction

Prediction generation

- infer latent actions $a_{1:T-1}$ from the observations $o_{1:T}$
- generate predictions $\hat{o}_{2:T}$ autoregressively

prediction accuracy averaged across environments



FLAM generates more accurate predictions than monolithic LAM and other baselines.

Results - prediction

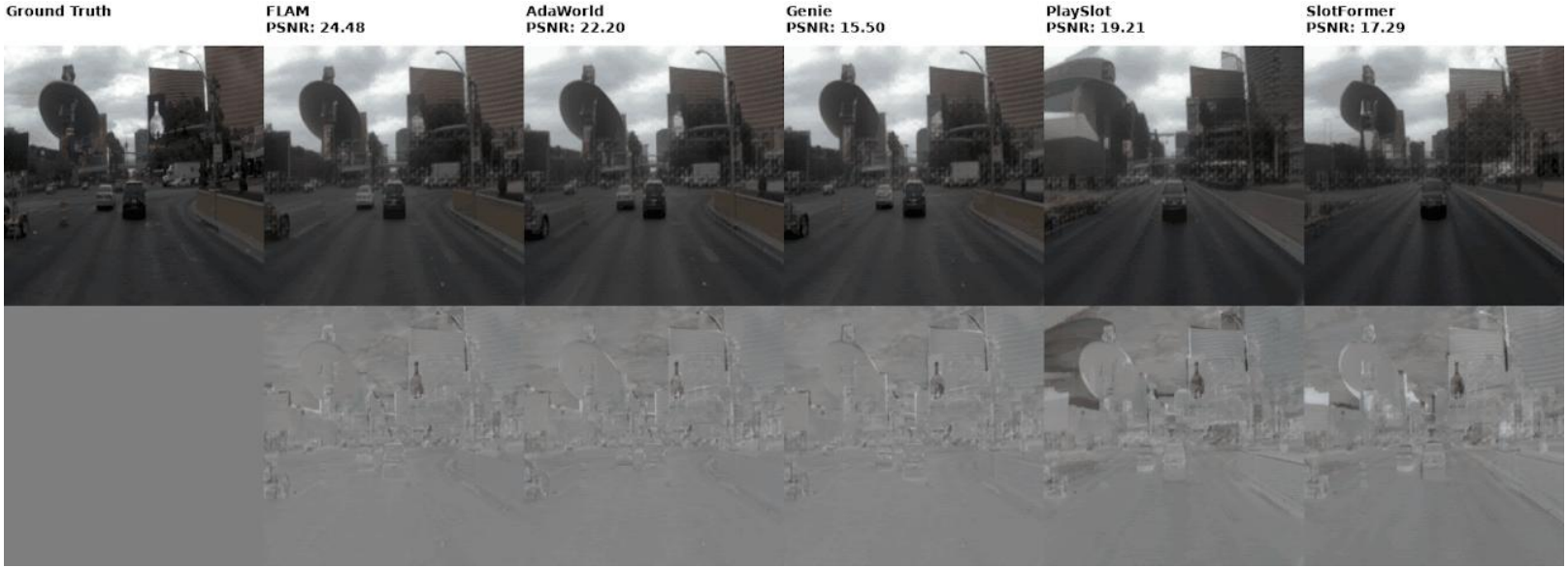
Prediction generation

- infer latent actions $a_{1:T-1}$ from the observations $o_{1:T}$
- generate predictions $\hat{o}_{2:T}$ autoregressively

FLAM generates more accurate predictions than monolithic LAM and other baselines.

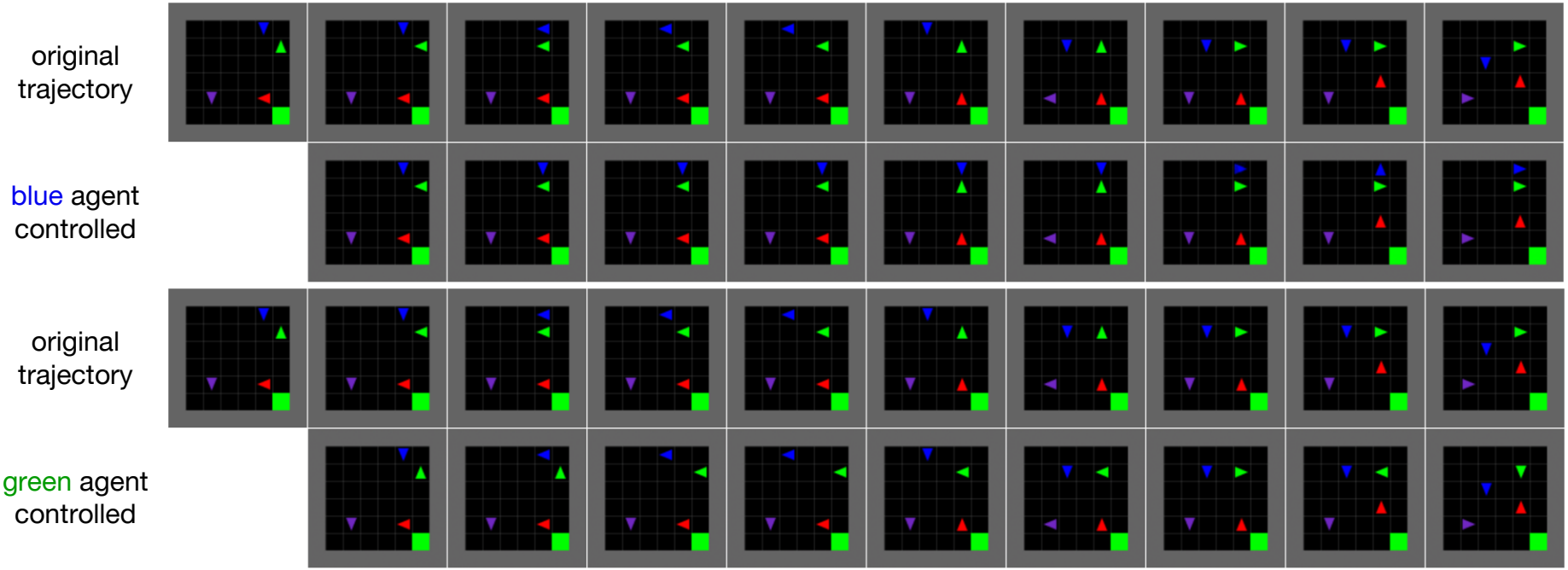
ground-truth / rollouts

prediction error



Results - controllability

Replace one factor's latent actions with random sampled latent actions.



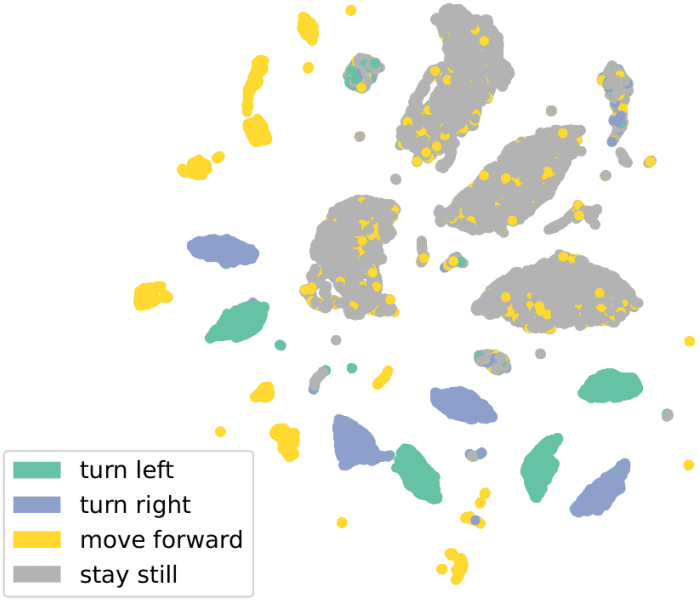
FLAM enables changing one factor without affecting others.

Monolithic LAM cannot do that due to its entangled action space.

Results - controllability

We visualize the 2D UMAP projection of learned latent actions.

- Points are colored by the ground-truth actions.
- The latent actions form well-separated clusters that align with the true actions.
- Overlap is expected: an agent may take a *move forward* action but remain stationary due to being blocked.

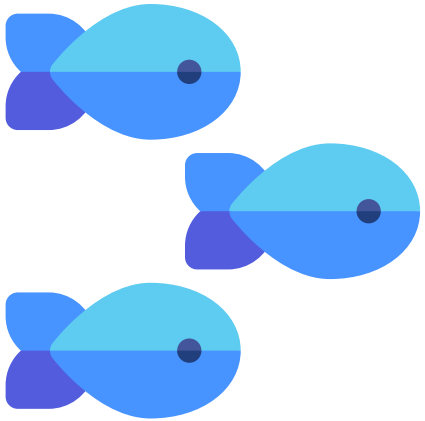


Results - factorization

The factorization learned by FLAM depend on

- the degree of action independence across state factors,
- the number of factors K.

When many factors have similar or correlated actions, a small K encourages the model to merge them into the same factor.

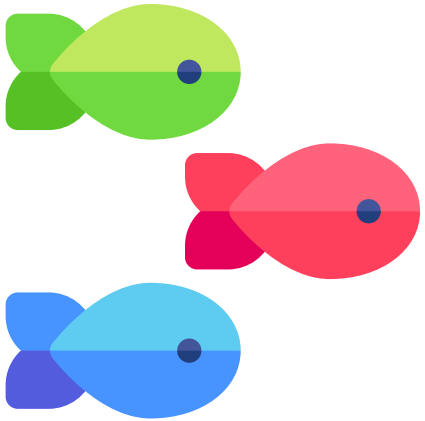


Results - factorization

The factorization learned by FLAM depend on

- the degree of action independence across state factors,
- the number of factors K.

When K is large or actions are rather independent, the model allocates separate factors to entities to capture fine-grained action differences.



Results - factorization

The factorization learned by FLAM depend on

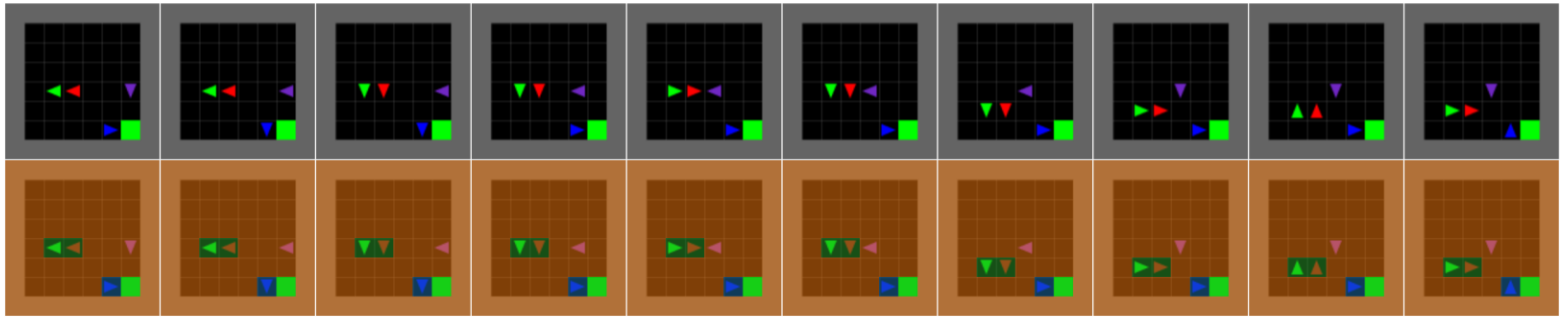
- the degree of action independence across state factors,
- the number of factors K.

When many factors have similar or correlated actions, a small K encourages the model to merge them into the same factor.

Green and red agents share the same actions.

ground truth

slot attention mask



They are grouped into the same factor.

Results - factorization

The factorization learned by FLAM depend on

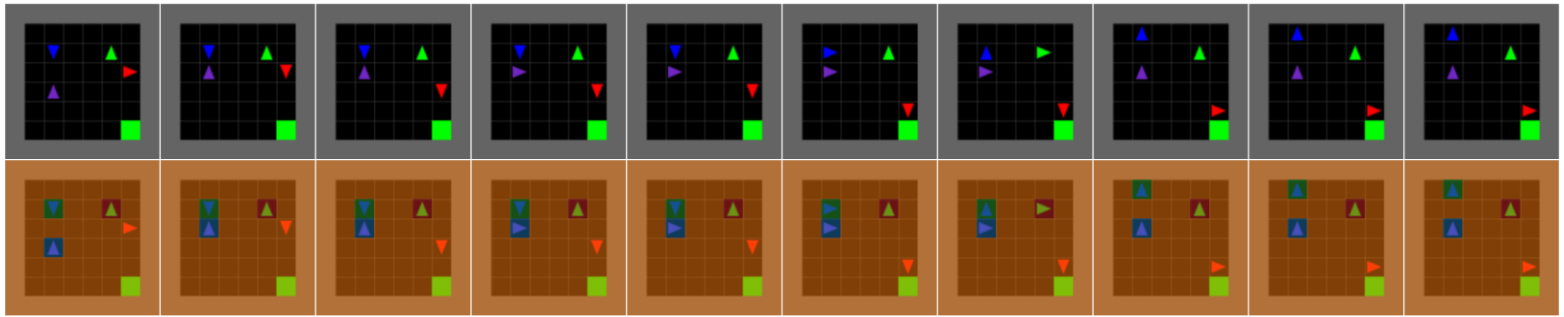
- the degree of action independence across state factors,
- the number of factors K.

When factors have independent actions or when K is large, the model allocates separate factors to agents to capture fine-grained action differences.

All agents act independently.

ground truth

slot attention mask



Each agent is allocated to a separate factor.

Contributions

State Abstractions

focus on task-relevant state factors

ICML 2022, AAAI 2024

Intrinsic Rewards

provide extra exploration signals

NeurIPS 2023

Unsupervised Skill Discovery

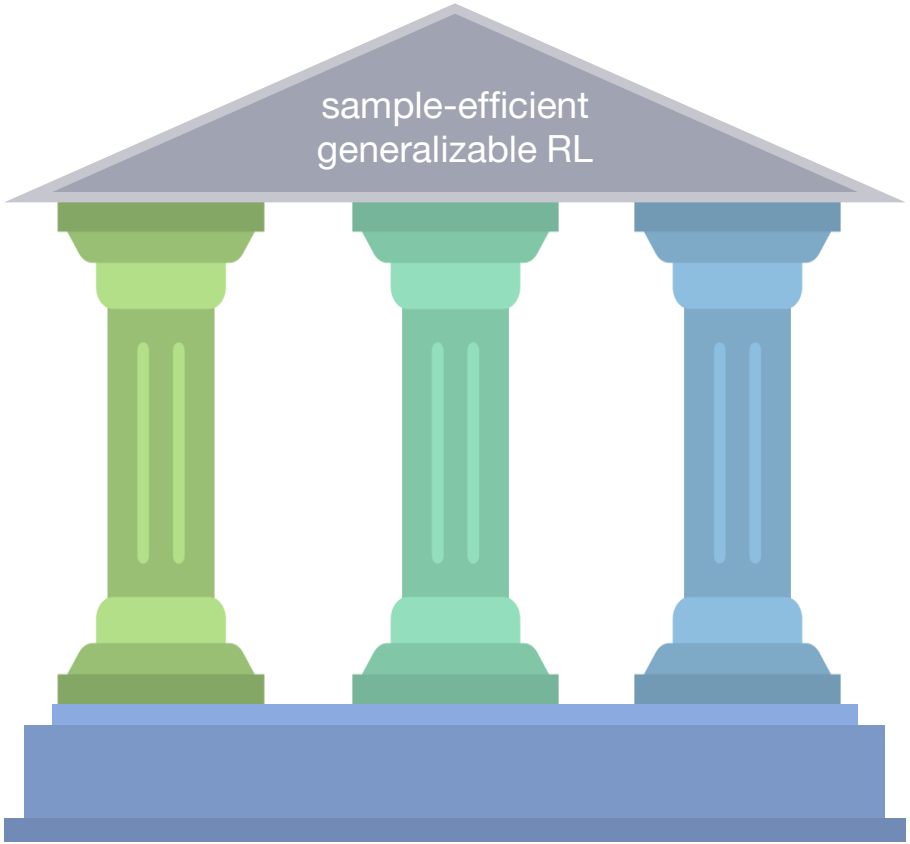
reduces exploration horizons

NeurIPS 2024

Representation Learning

extracts state factors from observations

NeurIPS 2025, preprint



Related work

RL (Bellman, 1957a,b; Howard, 1960; Lovejoy, 1991; Whittle, 1982; Bryson, 1996; Sutton, 1988; Watkins, 1989; Bertsekas and Tsitsiklis, 1996; Sutton, 1984, 1988; Mnih et al., 2015; Silver et al., 2016; Schulman et al., 2017; Wurman et al., 2022; Guo et al., 2025)

- sample-efficiency (Lin, 1992; Mnih et al., 2015; Silver et al., 2014; Haarnoja et al., 2018b; Fujimoto et al., 2018, Schaul et al., 2016; Brittain et al., 2019; Novati and Koumoutsakos, 2019; Zha et al., 2019; Sun et al., 2020; Zha et al., 2019; Fedus et al., 2020; Fujimoto et al., 2020; Sinha et al., 2022; Schmitt et al., 2020; Rolnick et al., 2019; Fernández and Veloso, 2006; Taylor et al., 2007; Barreto et al., 2017; Rusu et al., 2016a,b; Finn et al., 2017; Duan et al., 2016; Abbeel and Ng, 2004; Ramírez et al., 2022; Hester et al., 2018; Vecerík et al., 2017; Fujimoto et al., 2019; Kumar et al., 2019, 2020; Kostrikov et al., 2022)
- generalization (Huang et al., 2017; Kos and Song, 2017; Behzadan and Munir, 2017; Cobbe et al., 2019, 2020a; Tobin et al., 2017; Sadeghi and Levine, 2017; Peng et al., 2018; OpenAI et al., 2019; Rajeswaran et al., 2017; Laskin et al., 2020; Kostrikov et al., 2021; Igl et al., 2019; Raileanu et al., 2020; Wang et al., 2020b; Hansen and Wang, 2021; Amit et al., 2020; Vieillard et al., 2020; Hiraoka et al., 2022; Pinto et al., 2017; Rajeswaran et al., 2017; Pattanaik et al., 2018)

Causality (Wright, 1921; Fisher, 1935; Rubin, 1974; Spirtes et al., 1993; Pearl, 1995; Pearl and Bareinboim, 2014; Hernán and Robins, 2023; Peters et al., 2017)

- actual causality (Lewis, 1973; Mackie, 1974; Halpern and Pearl, 2005; Chockler and Halpern, 2004; Meliou et al., 2010; Halpern, 2016; Kueffner, 2021)
- causal discovery (Spirtes et al., 2000; Chickering, 2002; Peters et al., 2014; Perry et al., 2022; Zheng et al., 2018; Lachapelle et al., 2020; Zhu et al., 2020a; Wang et al., 2021; Lorch et al., 2021, 2022; Rolland et al., 2022; Gao et al., 2022; Reizinger et al., 2023)
- granger causality (Granger, 1969; Geweke, 1982; Eichler, 2007; Basu et al., 2015; Marinazzo et al., 2008; Ghysels et al., 2016; Heerah et al., 2021; Tank et al., 2022)
- causal representation learning (Chen et al., 2016; Higgins et al., 2017; Kumar et al., 2018; Chen et al., 2018; Kim and Mnih, 2018; Locatello et al., 2019; Khemakhem et al., 2020; Arjovsky et al., 2019; Brehmer et al., 2022; Ahuja et al., 2023; Lippe et al., 2023; Yang et al., 2021; Shen et al., 2022; Brehmer et al., 2022; Lippe et al., 2023; Deng et al., 2022)
- causal RL (Zhang et al., 2019a, 2020; Sontakke et al., 2021; Volodin et al., 2020; Tomar et al., 2021a; Fu et al., 2021a; Zhang et al., 2019b; Huang et al., 2022b; Buesing et al., 2019; Nair et al., 2019; Mozifian et al., 2020; Lyle et al., 2021; Seitzer et al., 2021; Lu et al., 2020; Sonar et al., 2021; Huang et al., 2022a; Pitis et al., 2020a; Liao et al., 2021; Li et al., 2020; Zhu et al., 2022)

Related work

Model-based RL (Sutton, 1991, 1990; Li and Todorov, 2004; Tedrake et al., 2010; Deisenroth and Rasmussen, 2011; Williams et al., 2017; Chua et al., 2018; Nagabandi et al., 2018; Li and Todorov, 2004; Tedrake et al., 2010; Deisenroth and Rasmussen, 2011; Du et al., 2019; Janner et al., 2022; Kurutach et al., 2018; Janner et al., 2019; Pitis et al., 2020b; Zhang et al., 2024a; Liu et al., 2024; Ha and Schmidhuber, 2018; Hafner et al., 2020, 2021, 2024; Micheli et al., 2023, 2024; Cohen et al., 2024; Nagabandi et al., 2018; Feinberg et al., 2018; Amos et al., 2021; Goyal et al., 2021b, 2020, 2021a; Mattner et al., 2023; Zhang et al., 2024b; Zhao et al., 2022; Sehgal et al., 2024; Baek et al., 2025)

State abstractions (Givan et al., 2003; Ravindran and Barto, 2003; Dean et al., 1997; Ferns et al., 2004; Boutilier et al., 2000a; Givan et al., 2003; Chapman and Kaelbling, 1991; McCallum, 1996a; Jong and Stone, 2005b; Abel et al., 2016; Gelada et al., 2019; Zhang et al., 2021b; Castro et al., 2021; Wang et al., 2022a; Huang et al., 2022c; Hansen-Estruch et al., 2022; Tomar et al., 2021b; Allen et al., 2021)

Intrinsic reward functions (Schmidhuber, 1991)

- curiosity (Brafman and Tennenholtz, 2002; Kearns and Singh, 2002; Bellemare et al., 2016; Ostrovski et al., 2017; Tang et al., 2017; Martin et al., 2017; Machado et al., 2018; Lee et al., 2019; Stadie et al., 2015; Pathak et al., 2017, 2019; Burda et al., 2019; Houthooft et al., 2016; Achiam and Sastry, 2017; Hester and Stone, 2013; Shyam et al., 2019)
- empowerment (Klyubin et al., 2005b; Salge et al., 2014; Mohamed and Rezende, 2015; Gregor et al., 2017; Levy et al., 2023; Levy, 2025; Zhao et al., 2021; Seitzer et al., 2021)

Unsupervised skill discovery

- mutual information (Gregor et al., 2017; Florensa et al., 2017; Achiam et al., 2018; Eysenbach et al., 2019; Lee et al., 2019; Sharma et al., 2020; Liu and Abbeel, 2021a; Zhang et al., 2021c; Campos et al., 2020; Kim et al., 2021; Laskin et al., 2022)
- state coverage (Liu and Abbeel, 2021b; Campos et al., 2020; Park et al., 2022, 2023a, 2024b)
- goal-conditioned RL (Warde-Farley et al., 2019; Nair et al., 2018a; Pong et al., 2020; Liu et al., 2025; Eysenbach et al., 2021,2022,2024; Myers et al., 2024, 2025; Durugkar et al., 2021; Durugkar, 2023; Ma et al., 2022; Agarwal et al., 2023)
- basis functions (Wu et al., 2018; Touati and Ollivier, 2021; Touati et al., 2022; Park et al., 2024a; Agarwal et al., 2024; Sikchi et al., 2024)

Related work

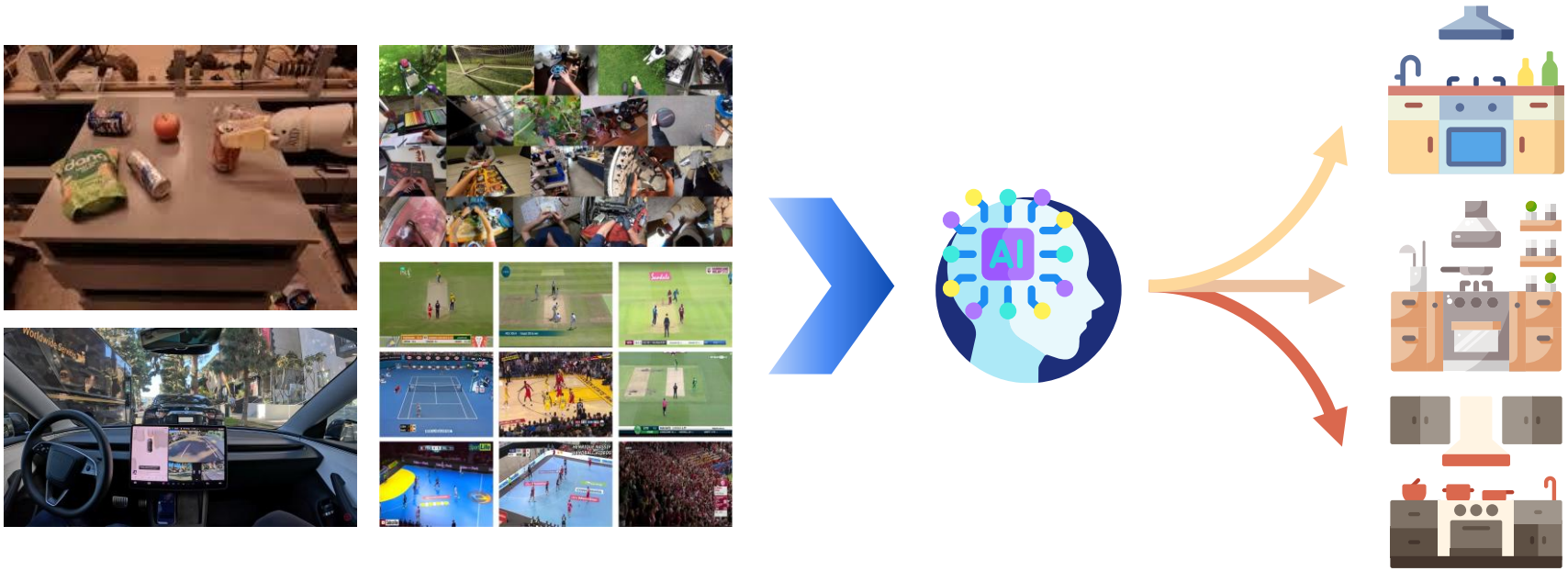
Object-centric representations (Burgess et al., 2019; Greff et al., 2019; Engelcke et al., 2020; Locatello et al., 2020; Seitzer et al., 2023; Wu et al., 2023a, 2023b; Jiang et al., 2023; Kipf et al., 2022; Elsayed et al., 2022; Aydemir et al., 2023; Seitzer et al., 2023; Zadaianchuk et al., 2023; Manasyan et al., 2024; Fan et al., 2024; Singh et al., 2024; Wu et al., 2023b; Singh et al., 2025; Akan and Yemez, 2025; Baek et al., 2025; Mosbach et al., 2025; Jiang et al., 2024; Yoon et al., 2023)

Latent action models (Edwards et al., 2019; Schmidt and Jiang, 2024; Bruce et al., 2024; Nikulin et al., 2025; Ye et al., 2024; Menapace et al., 2021; Zhang et al., 2022a; Ye et al., 2022, 2024; Baker et al., 2022; Chen et al., 2024; Villar-Corrales and Behnke, 2025; Klepach et al., 2025)

Future work

A foundation causal RL (world + policy) model that

- learns from diverse environments,
- enables sample-efficient adaptation to new environments.



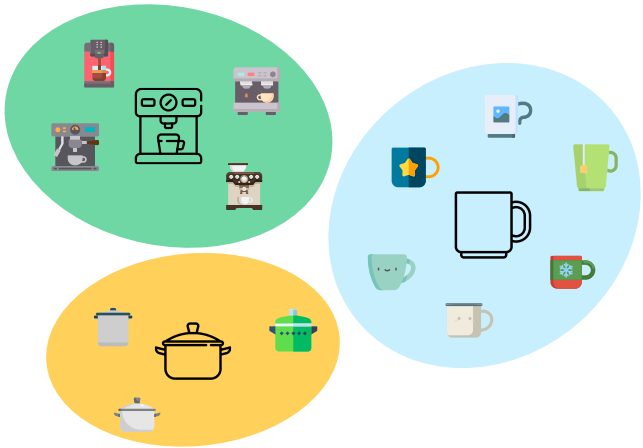
Future work

A foundation causal RL (world + policy) model that

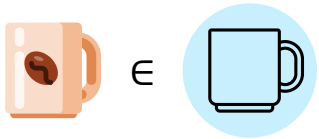
- learns from diverse environments,
- enables sample-efficient adaptation to new environments.

Open problems to address

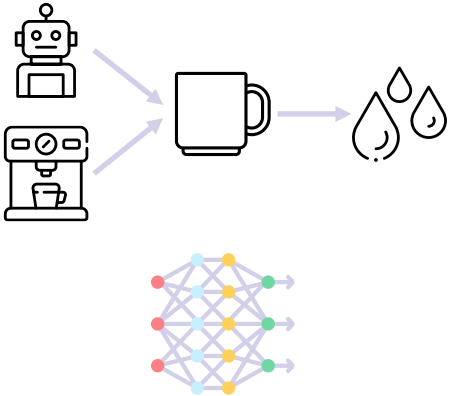
- reusable dynamics learning + efficient continual learning
  - identify entity types and reuse their learned causal relationships and dynamics models



entity types
learned during training



identify type of each entity
in a new environment



reuse causal relationships
and dynamics models

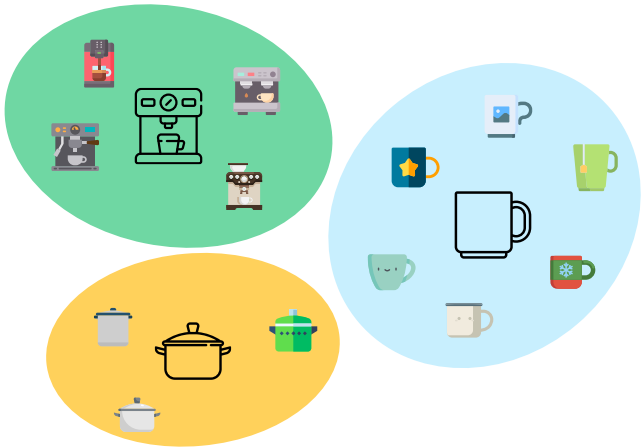
Future work

A foundation causal RL (world + policy) model that

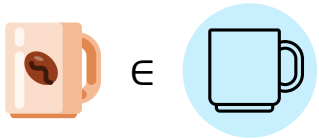
- learns from diverse environments,
- enables sample-efficient adaptation to new environments.

Open problems to address

- reusable dynamics learning + efficient continual learning
  - identify entity types and reuse their learned causal relationships and dynamics models



entity types
learned during training



identify type of each entity
in a new environment

reuse causal relationships
and dynamics models

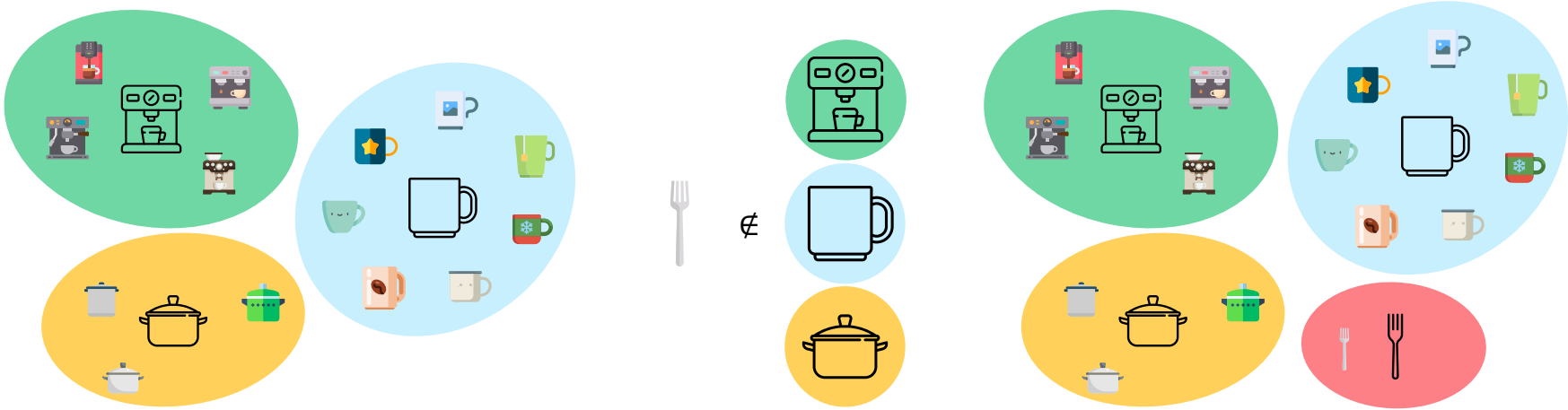
Future work

A foundation causal RL (world + policy) model that

- learns from diverse environments,
- enables sample-efficient adaptation to new environments.

Open problems to address

- reusable dynamics learning + efficient continual learning
  - identify entity types and reuse their learned causal relationships and dynamics models
  - identify unseen entities



Future work

A foundation causal RL (world + policy) model that

- learns from diverse environments,
- enables sample-efficient adaptation to new environments.

Open problems to address

- reusable dynamics learning + efficient continual learning
  - identify entity types and reuse their learned causal relationships and dynamics models
  - identify unseen entities
  - Identification based on visual appearance may not be reliable.
    - Sample-efficient interactions are necessary.

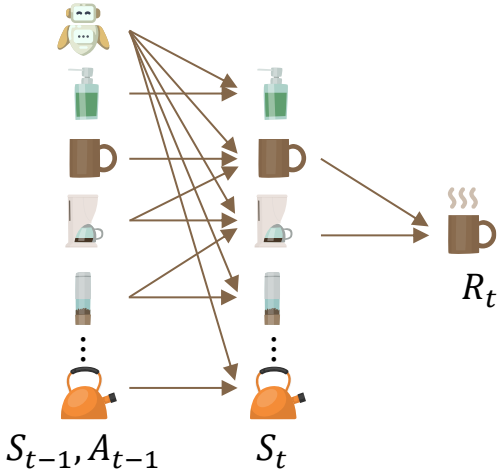
Future work

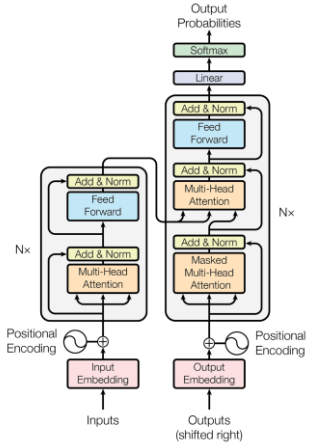
A foundation causal RL (world + policy) model that

- learns from diverse environments,
- enables sample-efficient adaptation to new environments.

Open problems to address

- reusable dynamics learning + efficient continual learning
- integration with existing foundation models (tokenization + architecture)





Thank you!








Thank you!