# Bisimulation Makes Analogies in Goal-Conditioned Reinforcement Learning

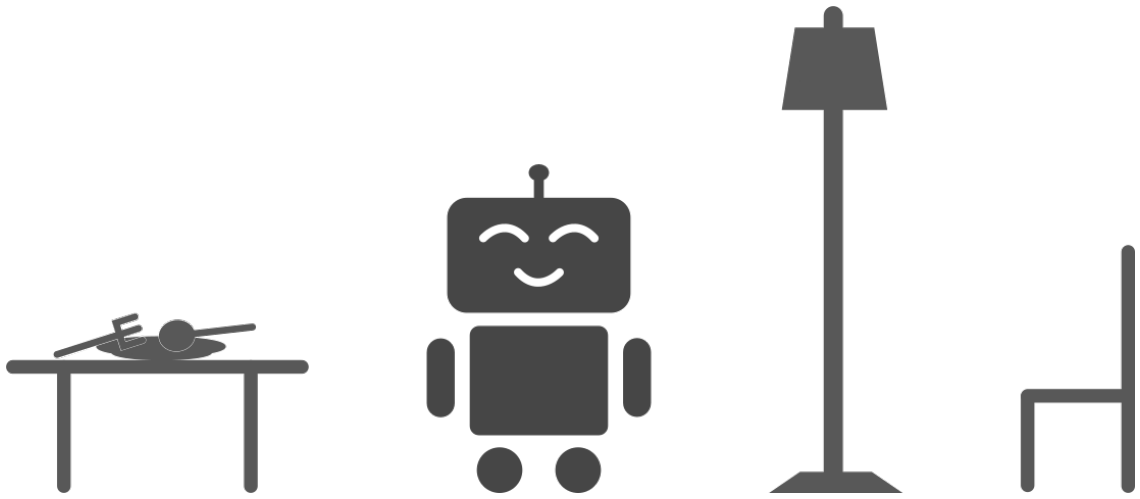
Philippe Hansen-Estruch [1], Amy Zhang [1 2], Ashvin Nair [1], Patrick Yin [1], Sergey Levine [1]

[1] University of California, Berkeley

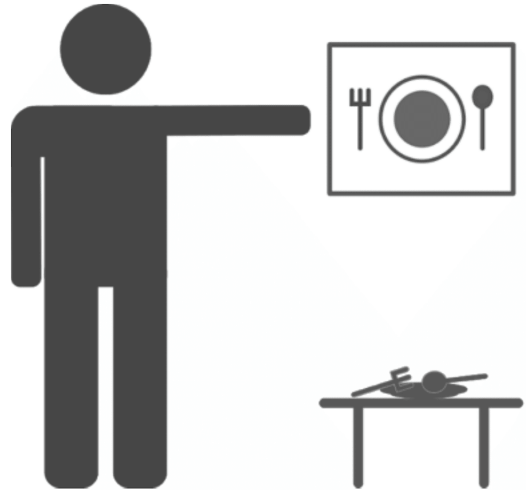
[2] Meta AI Research

# Goal-Conditioned Reinforcement Learning

- Learning generalizable multi-task agents is an important problem
  - Goal-Conditioned Reinforcement Learning (GCRL) $\pi(s, g)$
- A way to represent goals is needed
  - Prior work uses the exact goals



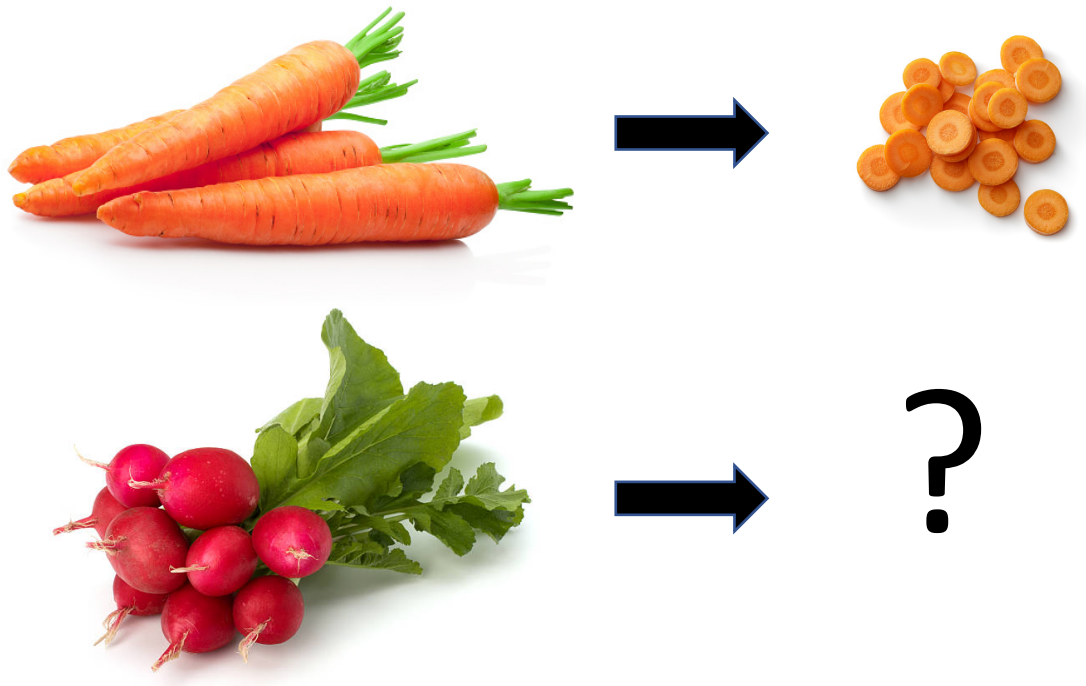
**Train: practice reaching various goals**



**Test: reach a specific goal**

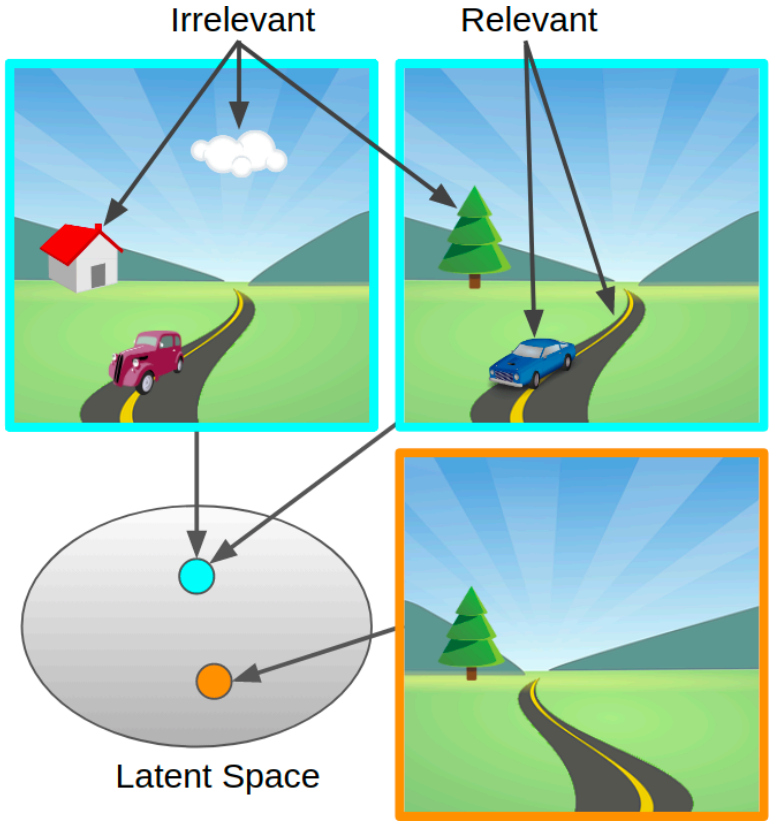
# What is the right way to represent goals?

- Conveying a task as a *transformation* in state is a more meaningful goal in RL than a single state
- We want to learn a task representation that captures *functional changes* in the environment
- This representation should retain information changeable by the agent and remain invariant to everything else
- Such a representation is broadly applicable to control



# A Functionally Equivariant Task Representation

- A transformation can be viewed as a state (start), goal (finish) pair
- Bisimulation has been shown to learn representations that capture functional invariance across states



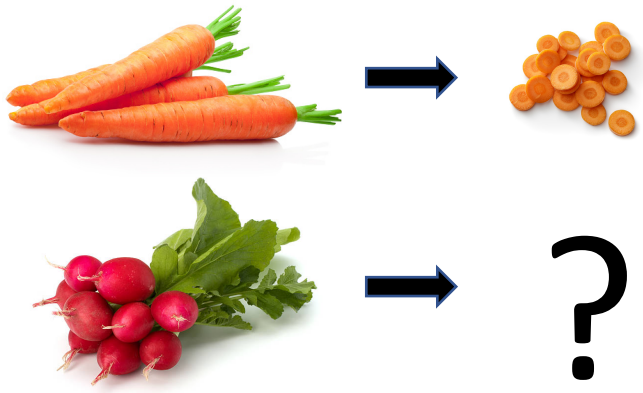

- Bisimulation could construct a task representation that captures functional equivariance across pairs

$$\phi(\text{🥕}, \text{🫐}) = \phi(\text{🌱}, \text{🧅})$$

# How do we use this task representation?

$$\phi(\,,\,) = \phi(\,,\,)$$

- We need a single state representation in order to compose new states with known existing tasks



Train:

$$\psi(\,) + \phi(\,,\,) = \psi(\,)$$

Test:

$$\psi(\,) + \phi(\,,\,) = \psi(\,)$$

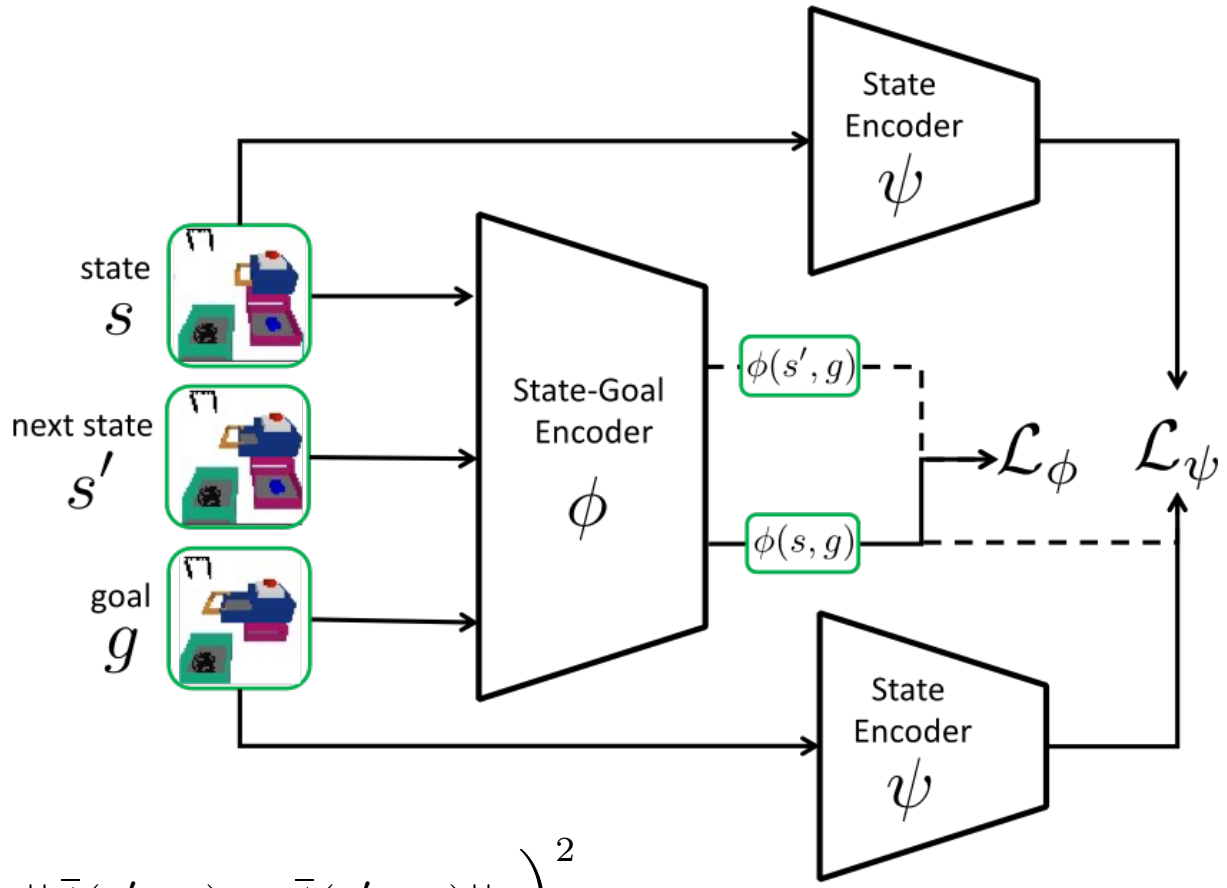
# Method

- We propose to learn *two* different representation spaces:
  1. A task embedding: a paired state representation that maps functionally equivalent tasks together.

$$\mathcal{L}_\phi = \left( ||\phi(s_i, g_i) - \phi(s_j, g_j)||_1 - ||r_i - r_j||_2 \quad -\gamma||\bar{\phi}(s'_i, g_i) - \bar{\phi}(s'_j, g_j)||_2 \right)^2$$

  2. A single state embedding: a single state representation capable of composing states and task embeddings to find the new goal.

$$\mathcal{L}_\psi = \left( (\bar{\phi}(s_i, g_i) - \bar{\phi}(g_i, g_i)) - (\psi(g_i) - \psi(s_i)) \right)^2$$



# Combining GC Bisimulation with Policy Learning

- Learn the representations while training the goal-conditioned policy in $\psi$ space.
- Using standard Offline RL, Policy trains on $\pi(\psi(s), \phi(s, g))$ and receives $\pi(\psi(s), \phi(s_a, g_a))$ during eval

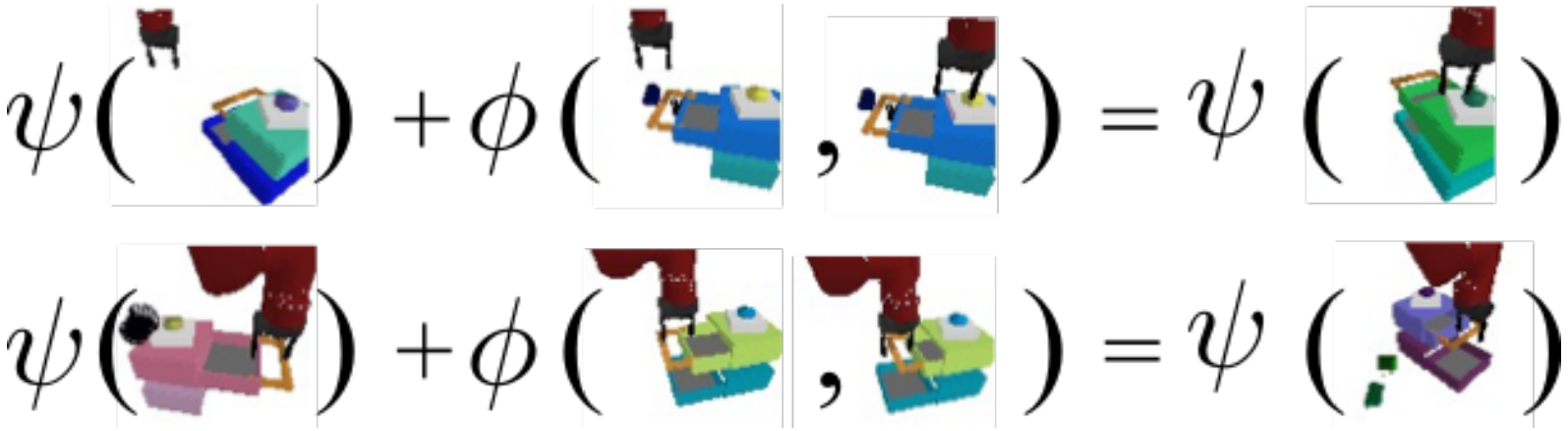
# Manipulation Experiments: Analogies Visualized

Start

Analogous State, Goal

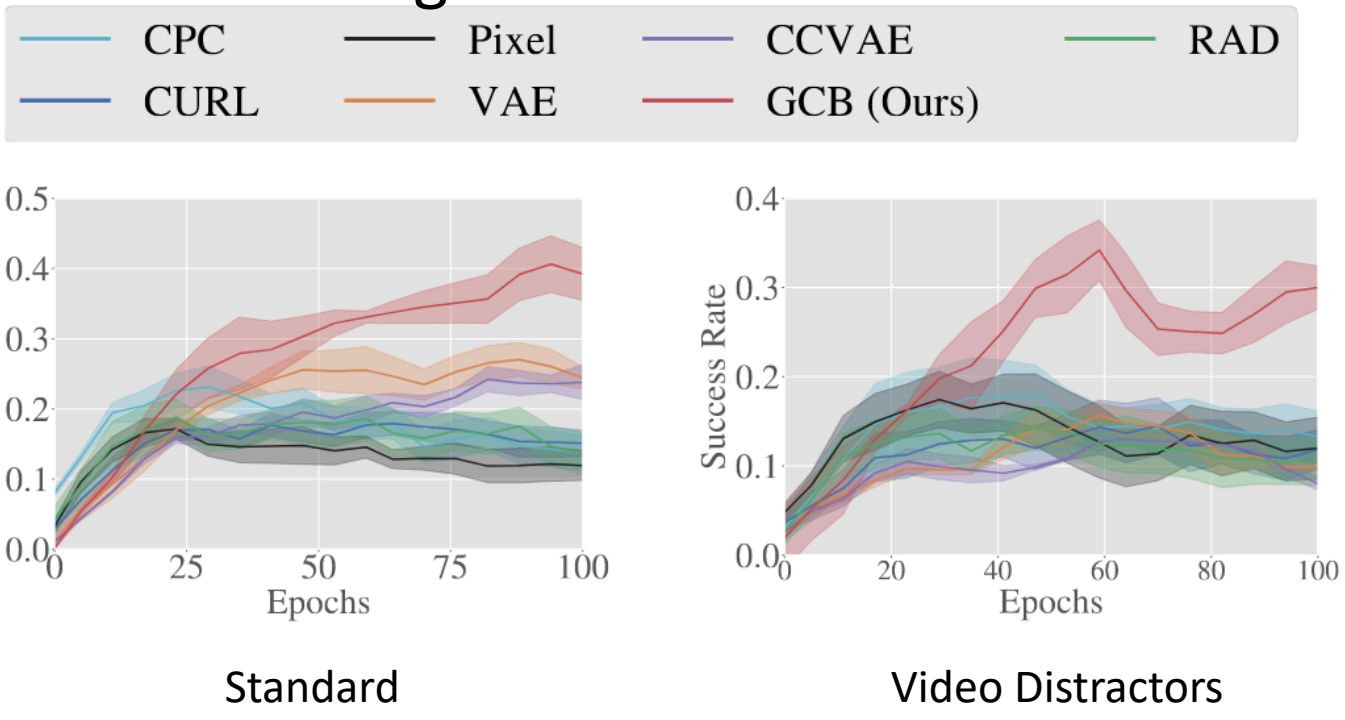
Implied Goal (1-NN in Dataset)

Video Distractor



# Manipulation Experiments: Using analogies for control

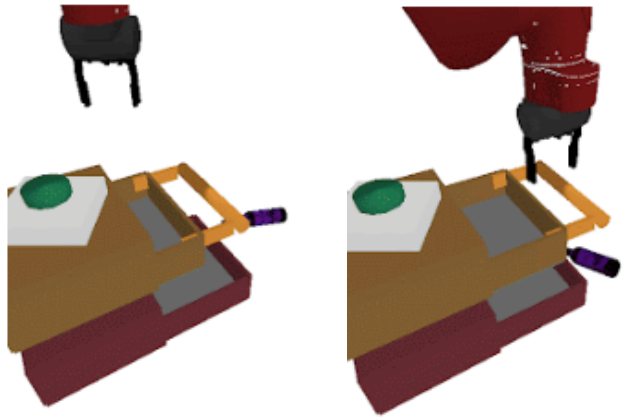
- If the agent is only given an example state-goal pair denoting a desired task, and a new state --- can it infer the new goal?



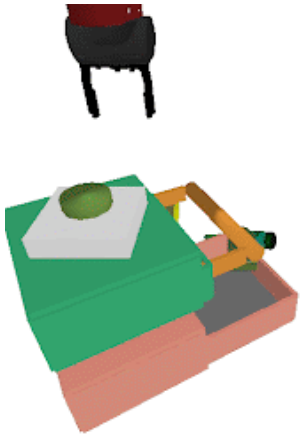

Standard

Video Distractors

State-goal pair:

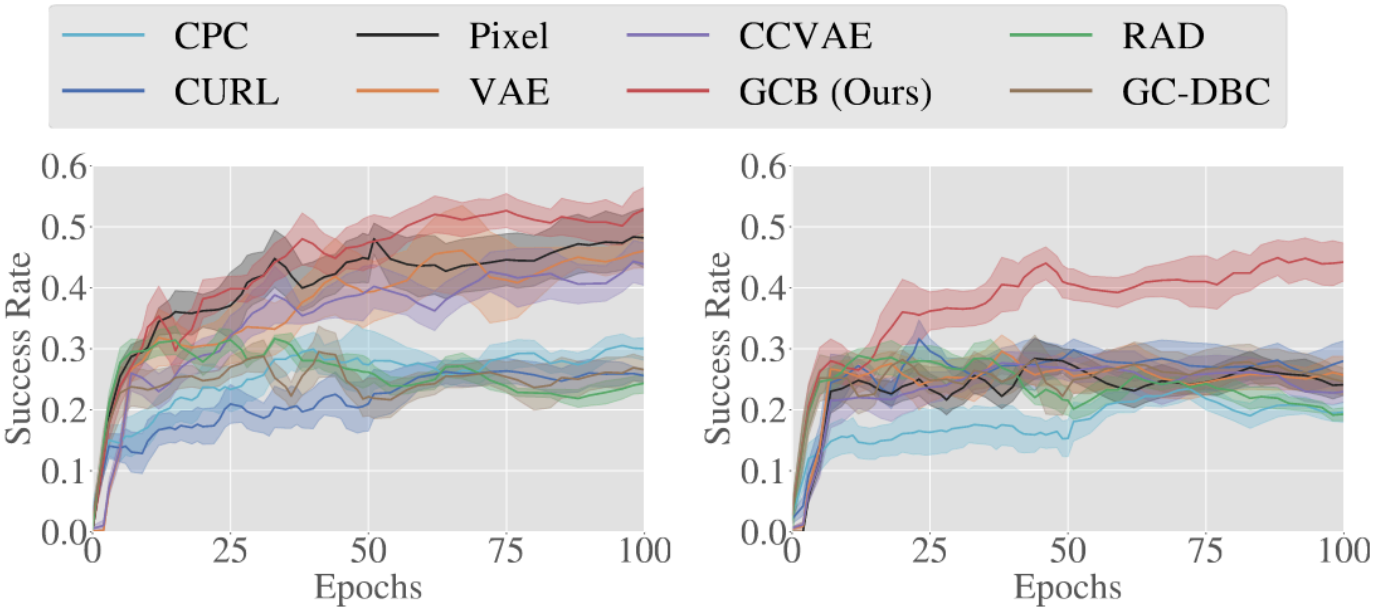


Rollout:



# Manipulation Experiments: Evaluating the standard goal-conditioned paradigm

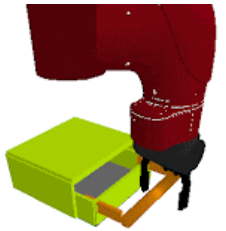
- Can our $\psi$ representation also lead to improved performance in the standard goal-conditioned paradigm?




Standard

Video Distractors

State-goal pair:



Rollout:

# Conclusion

- An ideal representation for GCRL captures *functional equivariance* which can be learned using Bisimulation

$$\phi(\text{  }, \text{  }) = \phi(\text{  }, \text{  })$$

- Coupled with a single state encoder, Goal-Conditioned Bisimulation is able to command goals with analogies

$$\psi(\text{  }) + \phi(\text{  }, \text{  }) = \psi(\text{  })$$

- Goal-Conditioned Bisimulation is also able to achieve SOTA performance on standard goal conditioned tasks