Value Function Spaces
Skill-Centric State Abstractions for Long-Horizon Reasoning
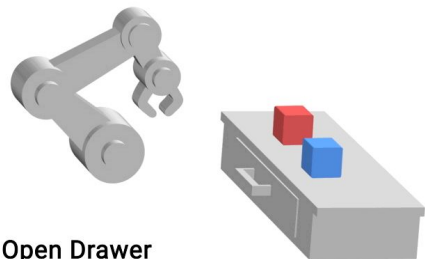
Dhruv Shah🐻, Peng Xu, Yao Lu, Ted Xiao,
Alexander Toshev, Sergey Levine🐻, Brian Ichter

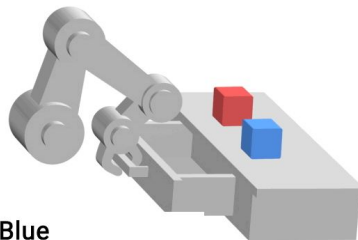International Conference on Learning
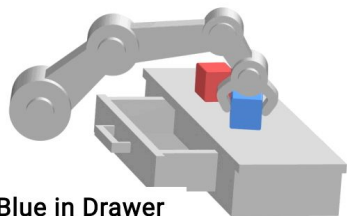Representations (ICLR), 2022
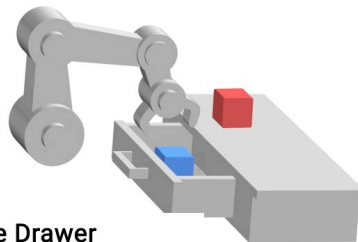
Google Research

🐻 also at UC Berkeley

Motivation

Open Drawer

Pick Blue

Place Blue in Drawer

Close Drawer

Google Research

Motivation

- Performing such tasks in real-world situations is challenging

    - Reasoning over long horizons

    - Parse high-dimensional observations

- How do we get state and action abstractions?

    - State abstraction should depend on the capabilities of the lower-level skills

Google Research

From Affordances to Skill Value Functions

- The notion of "values" in RL is closely related to affordances

  - Example, for sparse terminal reward: $V(s) = E(\gamma^t r(s))$

  - In general, value denotes the probability of success

- For any policy (or skill), its value function captures:

  - *preconditions* or *affordances* of the scene, i.e. where the skill can be used

  - *outcome,* i.e. did the skill execute successfully

- Given a finite set of skills, define a representation $Z(s_t)$ by stacking the corresponding skill value functions $Z(s_t) = [V_1(s_t), V_2(s_t), ..., V_k(s_t)]$
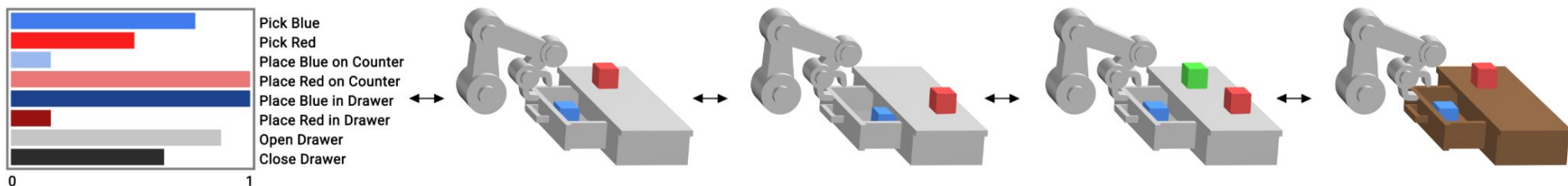
A Toy Example

"Place Blue in Drawer"

A Toy Example
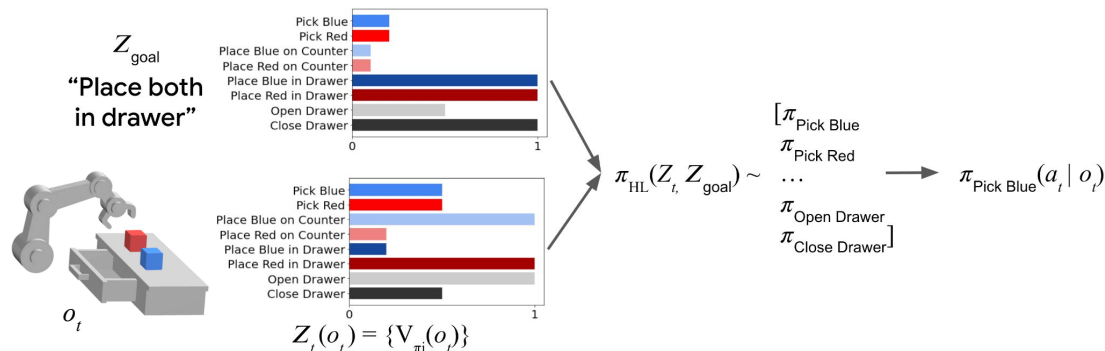
"Place Blue in Drawer"

A Toy Example

- The representation is robust to exogenous factors of variation
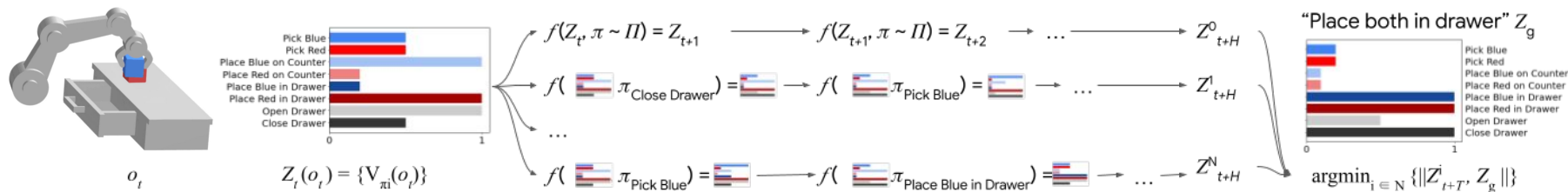 - generalization for free, modulo the underlying skills.

Instantiating RL Pipelines with VFS

- **<u>Model-Free</u>**: HRL with high-level policy over skills using VFS as state

- **<u>Model-Based</u>**: Random shooting using predictive model on VFS states

Instantiating RL Pipelines with VFS

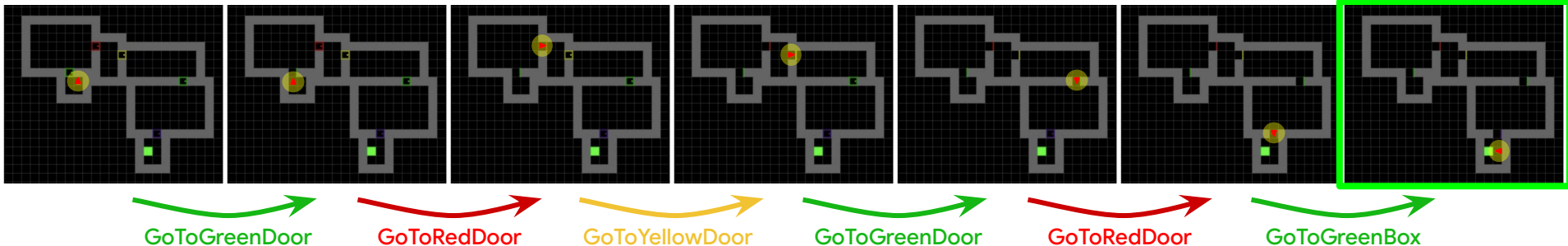Primitives — GoToGreenDoor GoToRedDoor GoToYellowDoor GoToPurpleDoor GoToGreenBox GoToRedBox GoToYellowBox GoToPurpleBox

Goal: GreenBox
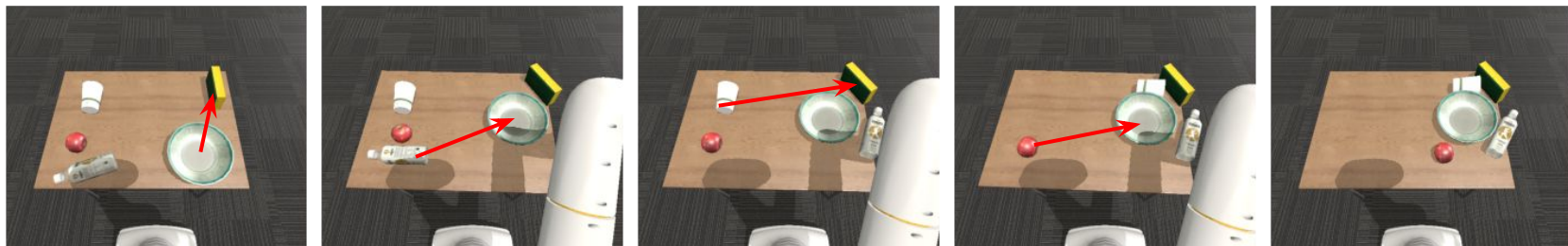

GoToGreenDoor → GoToRedDoor → GoToYellowDoor → GoToGreenDoor → GoToRedDoor → GoToGreenBox

Instantiating RL Pipelines with VFS

Primitives   Move "A" Near "B"

Goal: "Move All to Top Right Corner"

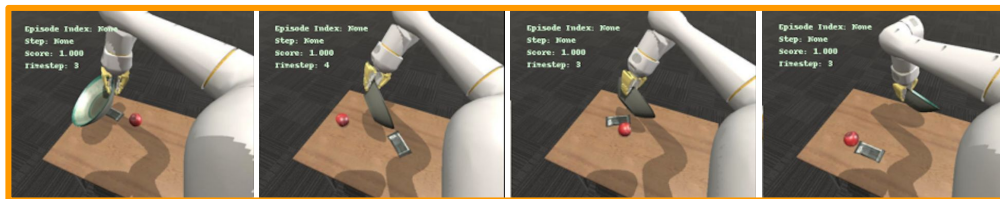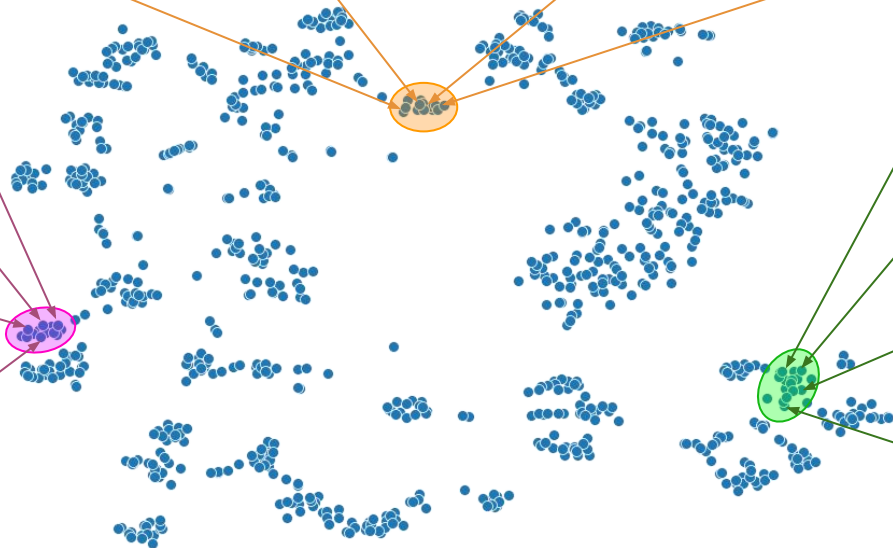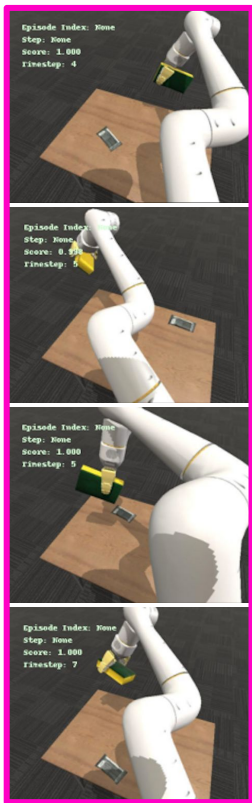Move Bowl Near Sponge   Move Bottle Near Bowl   Move Glass Near Sponge   Move Apple Near Bowl

Visualizing the Representations

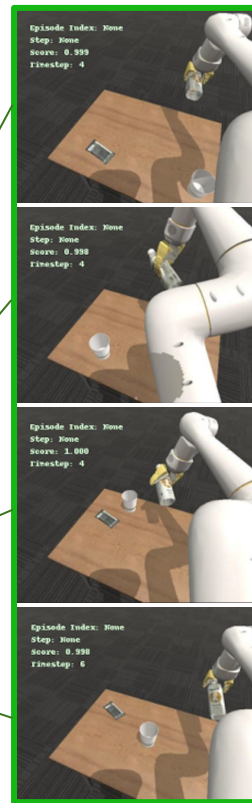bowl in hand

sponge in hand

bottle in hand

Value Function Spaces
Skill-Centric State Abstractions for Long-Horizon Reasoning

Dhruv Shah🐻, Peng Xu, Yao Lu, Ted Xiao,
Alexander Toshev, Sergey Levine🐻, Brian Ichter

🌐 cs.berkeley.edu/~shah/VFS
📫 shah@cs.berkeley.edu

Google Research

🐻 also at UC Berkeley