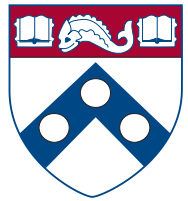


Specification-Guided Safety for Vision-Based Robot Learners

Dinesh Jayaraman

Assistant Professor, CIS, Upenn

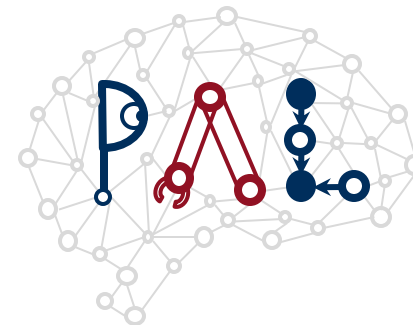
PIs: Rajeev Alur, Osbert Bastani, Eric Wong



Penn
Engineering

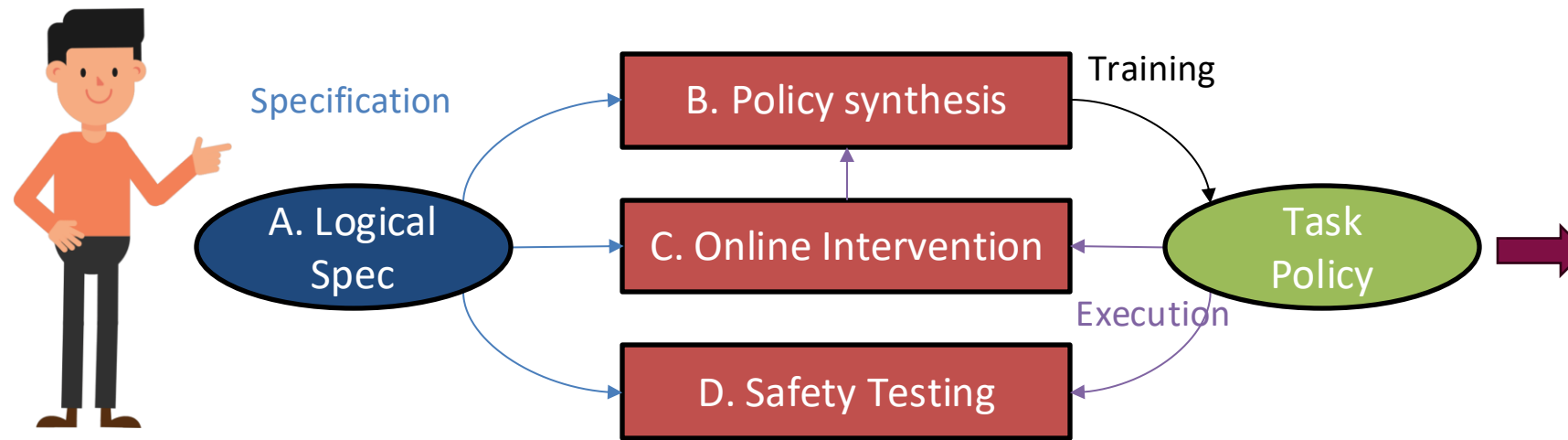
GRASP
Laboratory

General Robotics, Automation, Sensing & Perception Lab

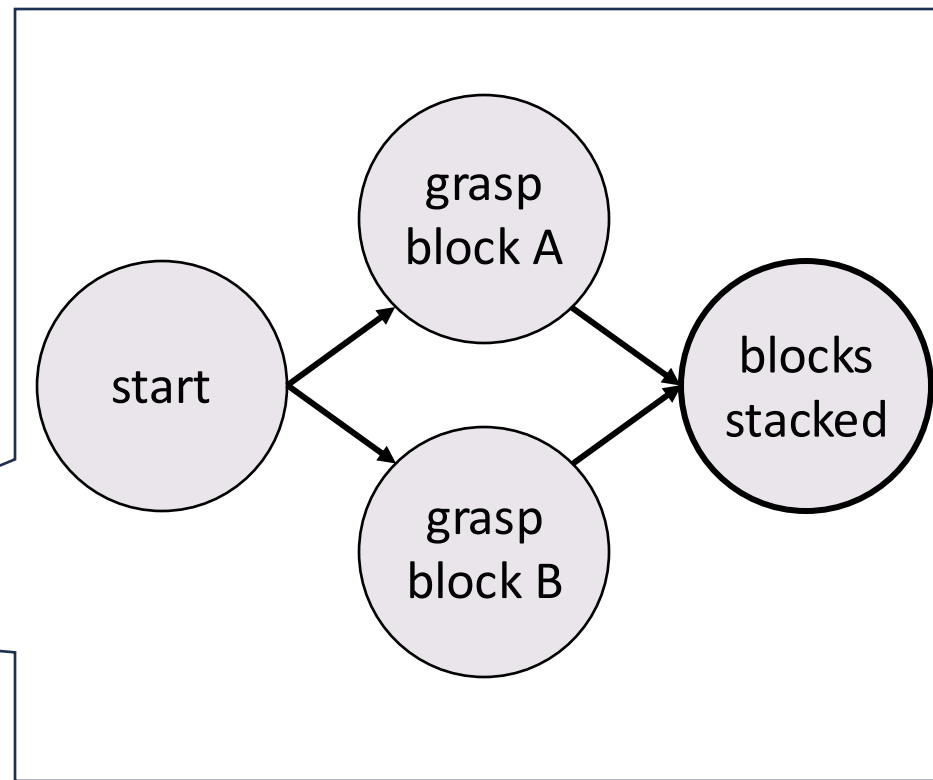
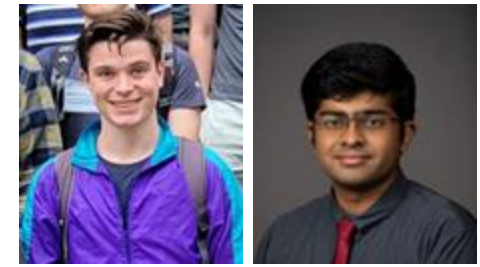


Perception,
Action, &
Learning Group

SPECSRL: Specification-guided Perception-enabled Conformal Safe Reinforcement Learning

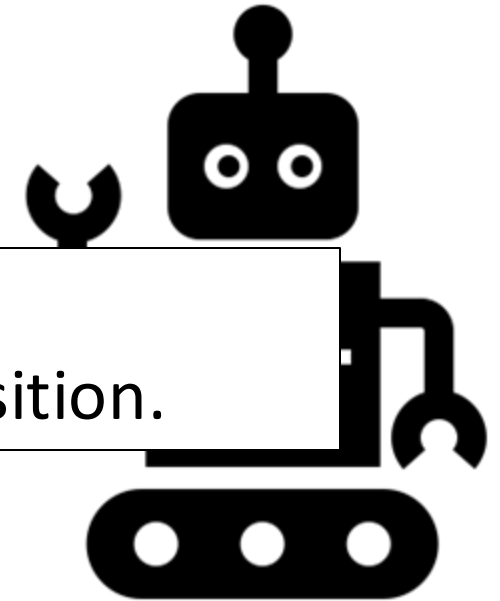


Landmark Graph

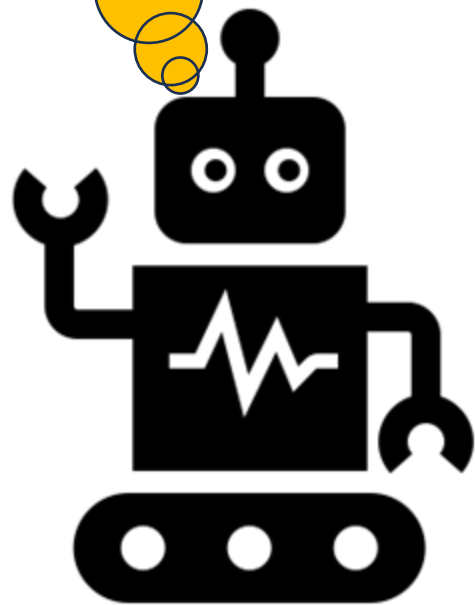
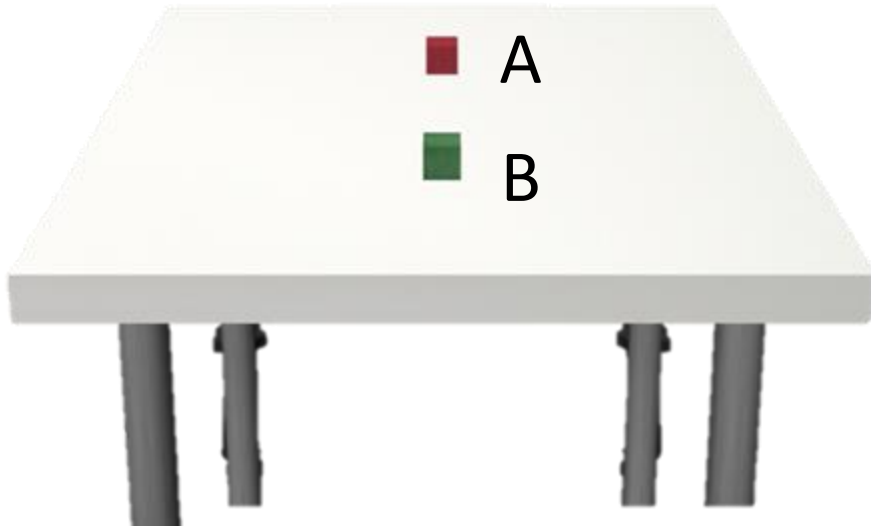
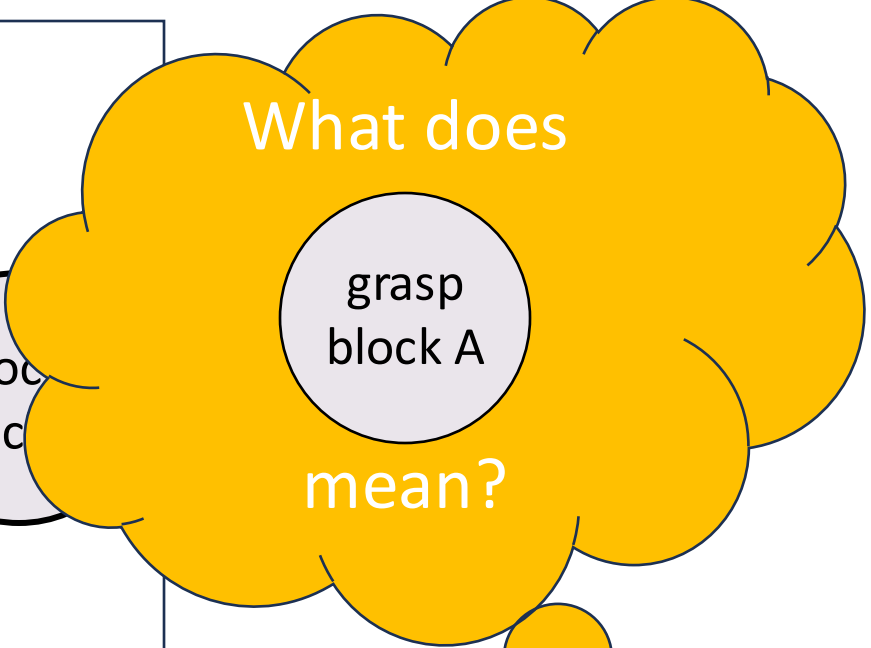
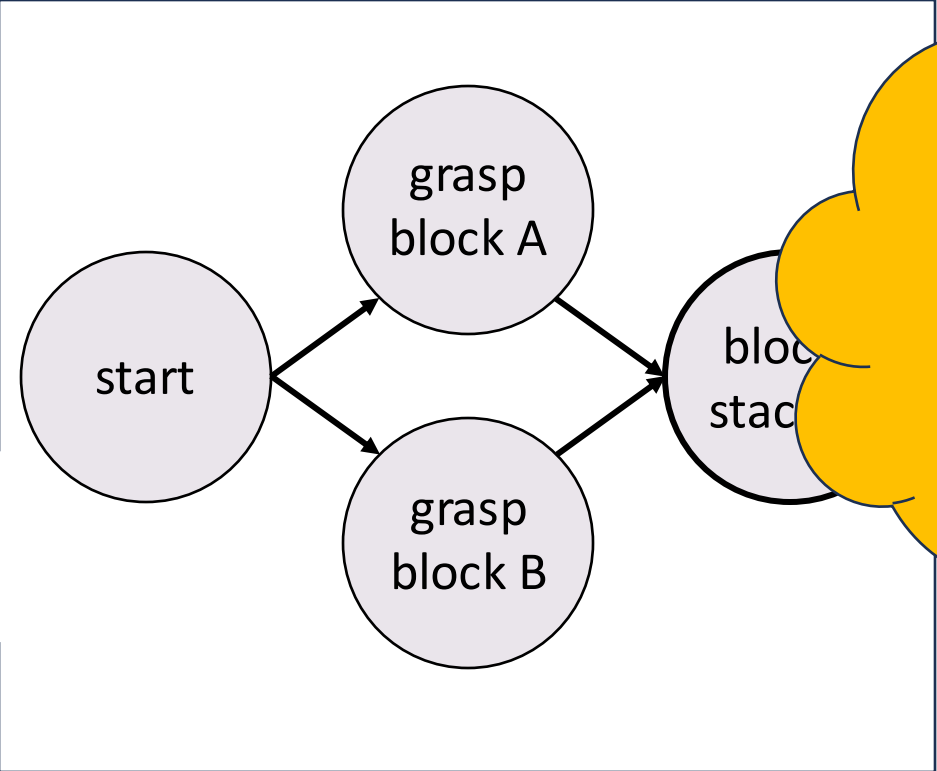


■ A

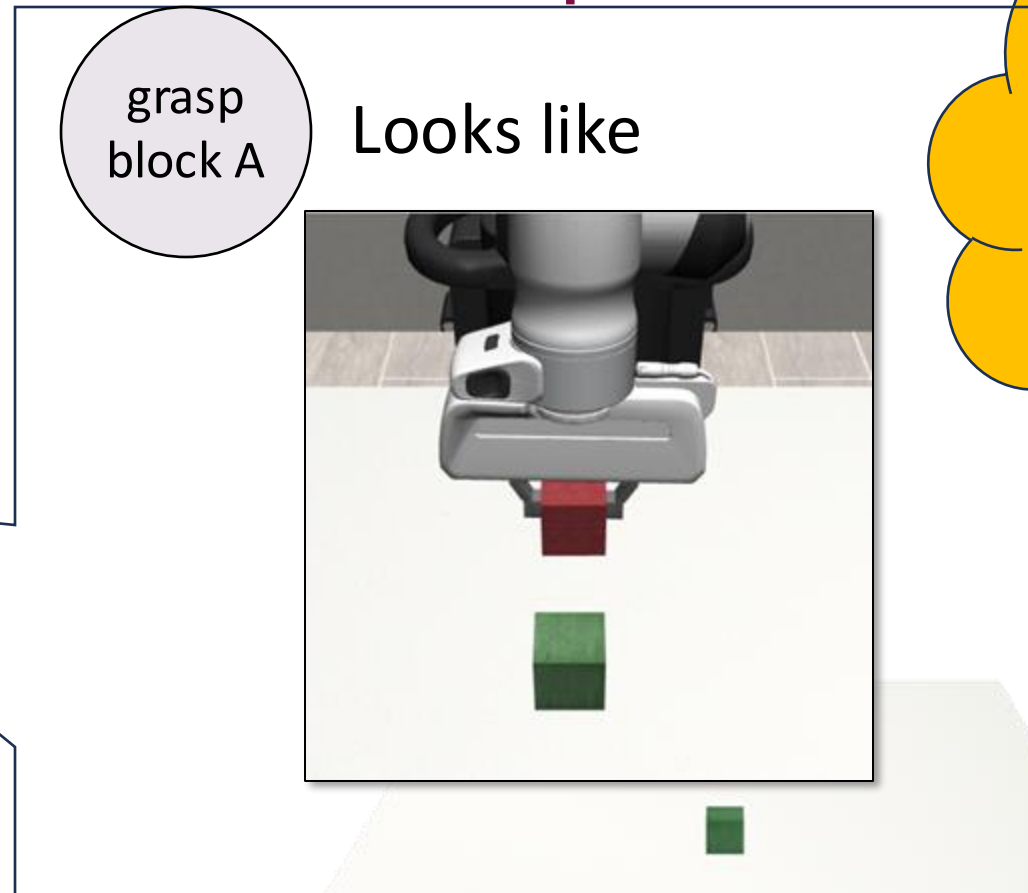
- Each node is a *landmark*, a set of states.
- Each edge represents a potentially achievable transition.



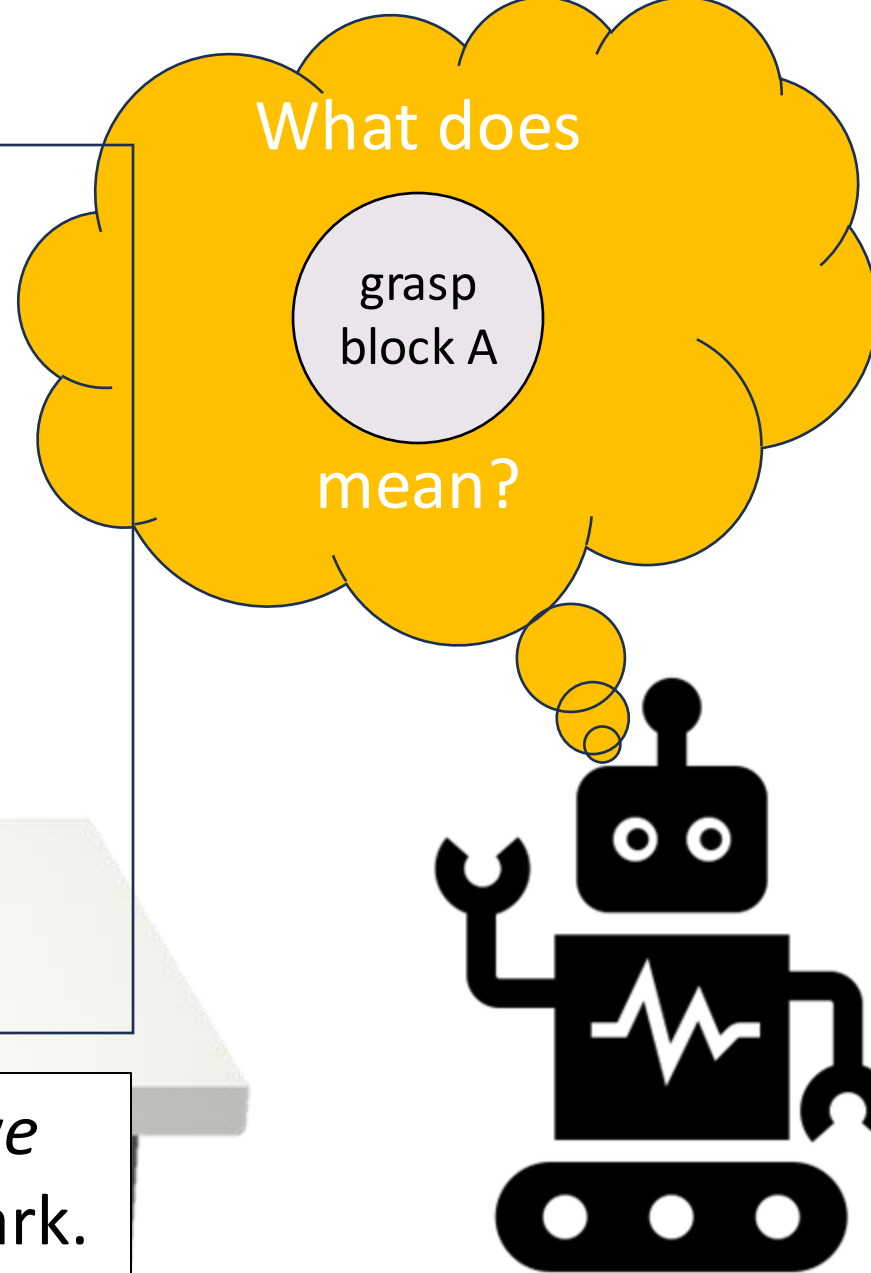
Landmark Graph

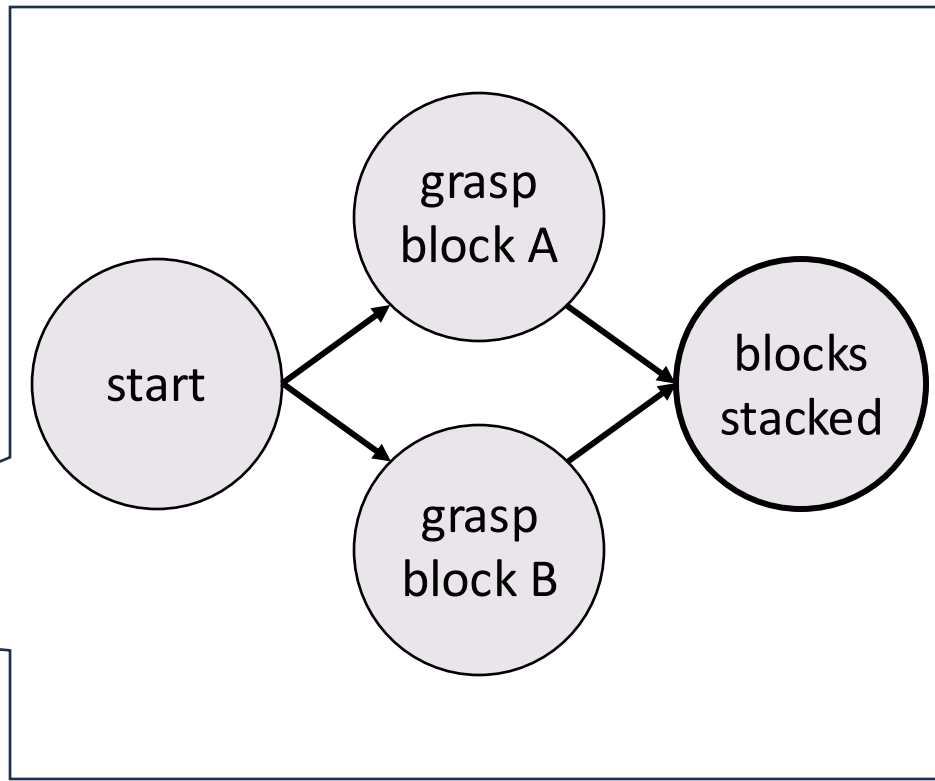


Illustrated Landmark Graph

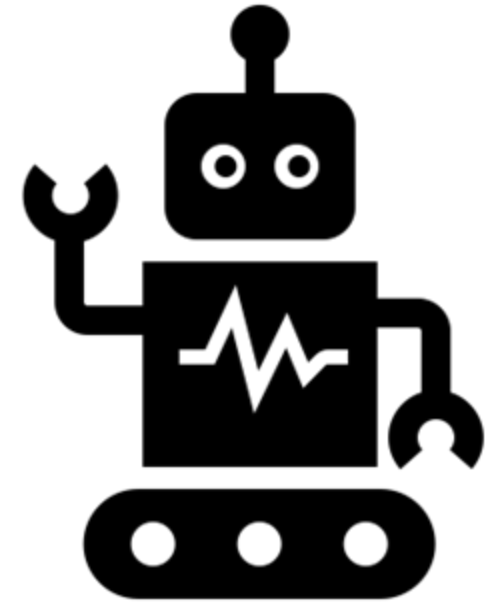


Teacher provides *illustrative observations* of each landmark.

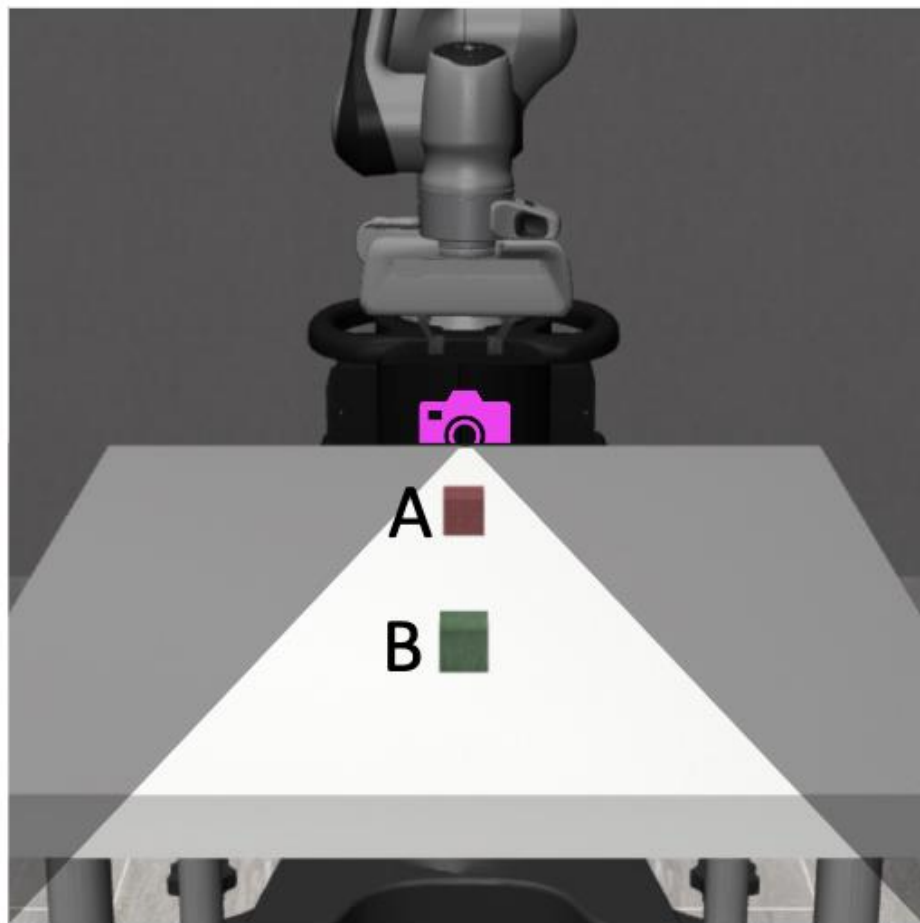




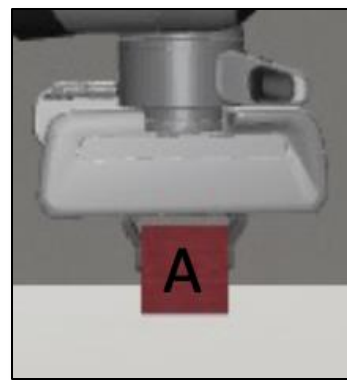
Which path to take?



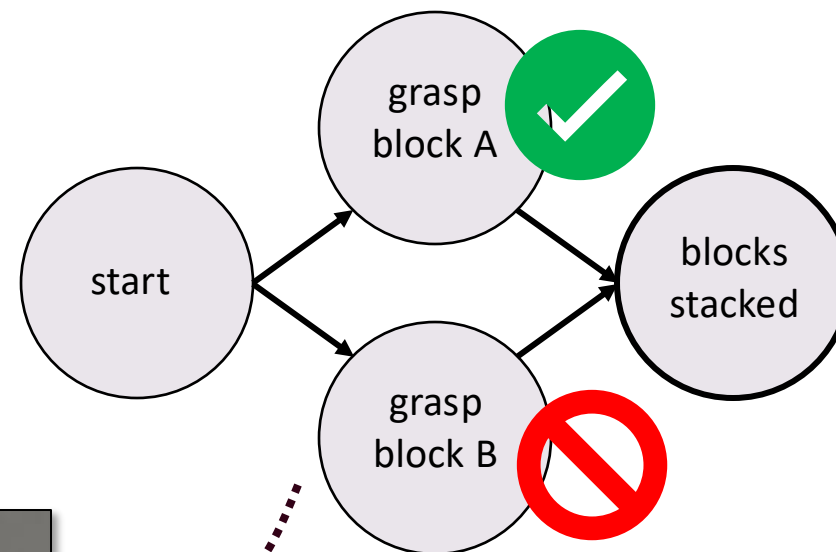
Flexibility to Choose Your Own Path



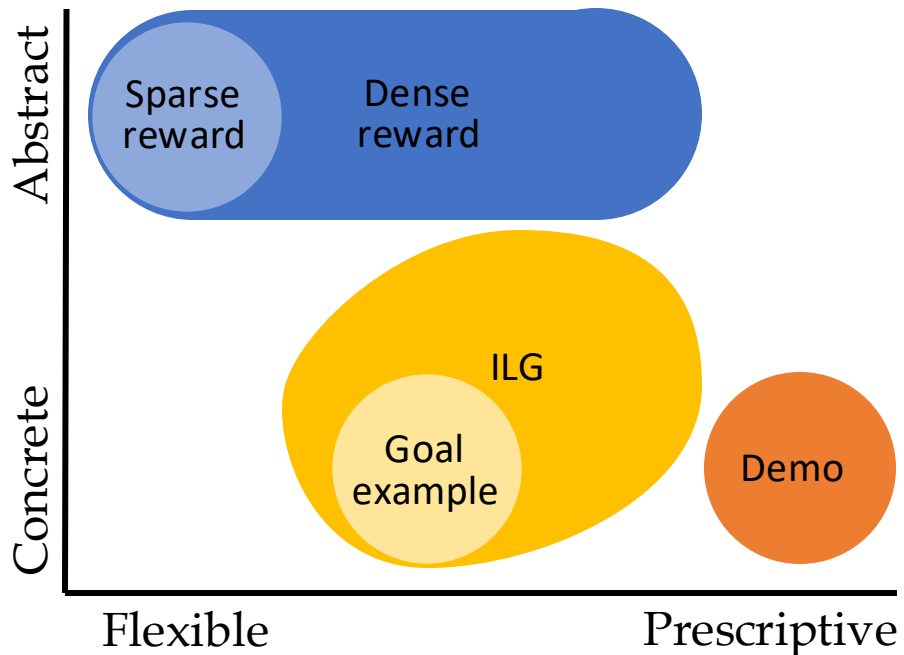
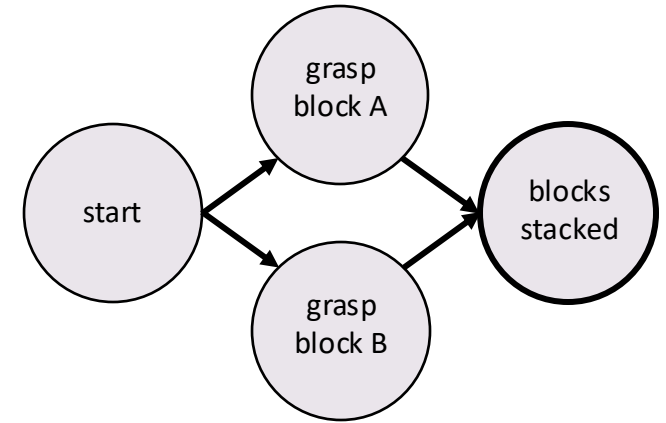
Initial layout



Ambiguous
camera image



ILGs Are Versatile And Expressive

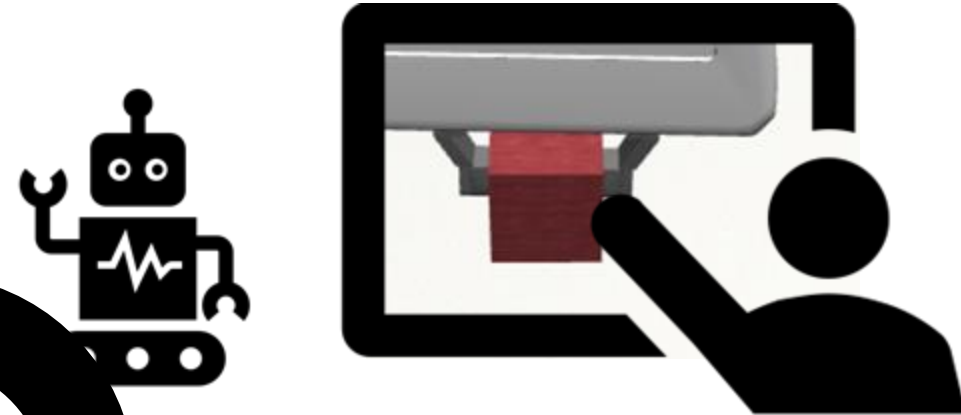
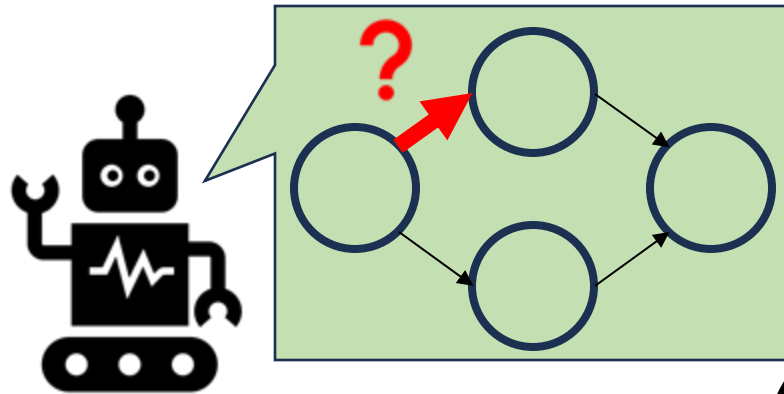


Teacher can select branching factors, and density of task specifications.

Controls the agent's freedom to select:

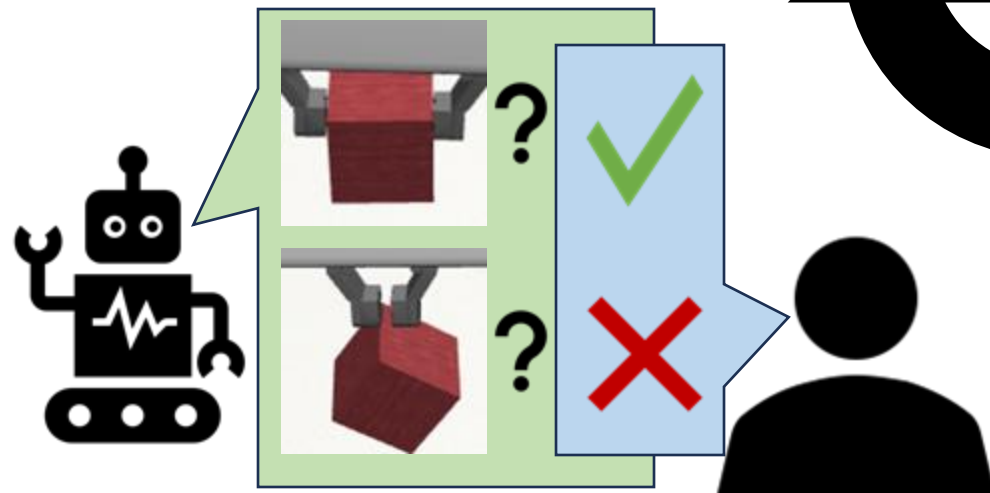
- High-level task strategy (path through the graph), or
- Low-level control to transition between connected nodes

An Interactive Learning Algorithm: ILG-Learn

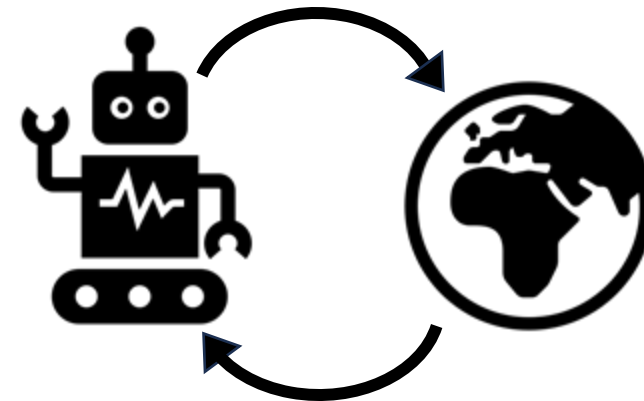


1. The learner selects an edge and requests illustrative observations to guide learning.

2. A human teacher provides illustrative observations of the destination landmark.

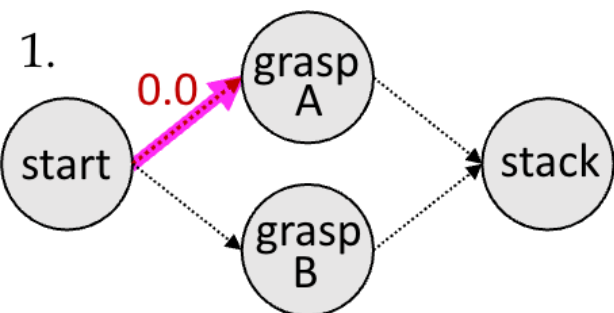


4. The learner executes the edge policy and queries the teacher for success/failure feedback.



3. The learner learns an *edge policy* from environmental interaction using example-based control.

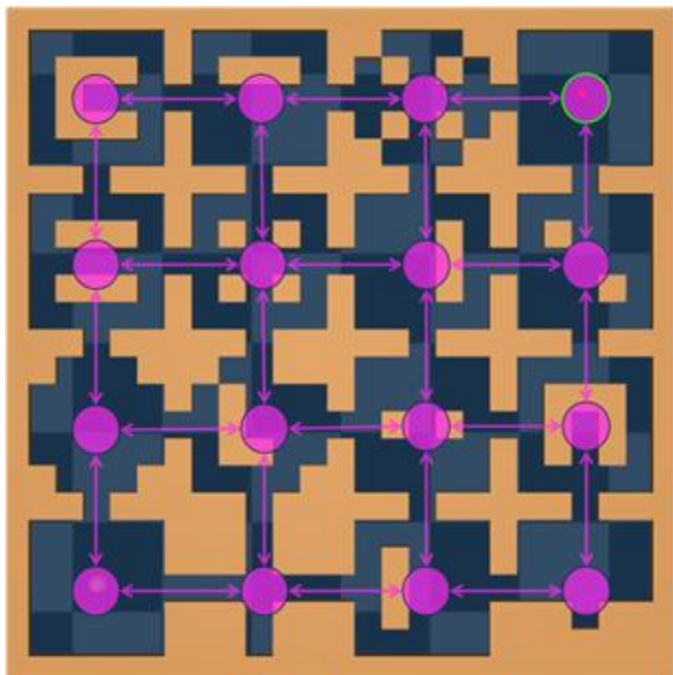
ILG-Learn For Block Stacking



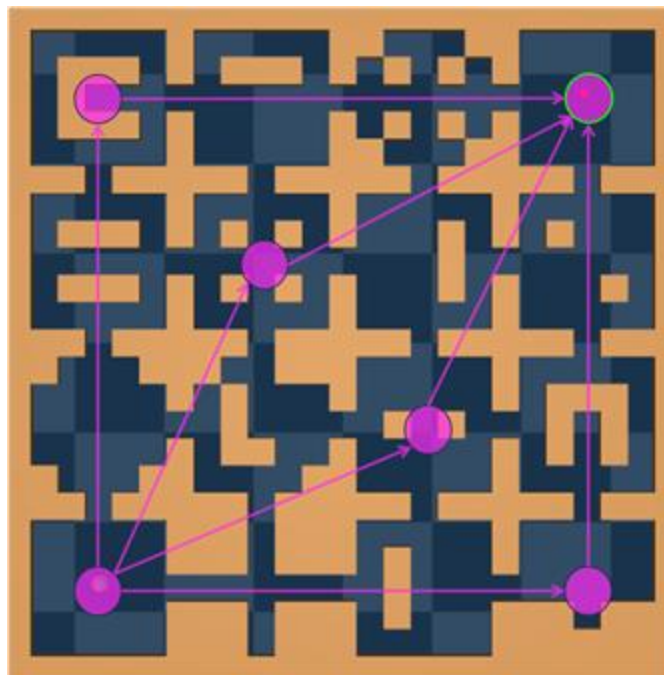
Selects a path to suit agent capabilities,
Learns a sequence of low-level policies to execute it.

Maze Suite For Thorough Evaluations

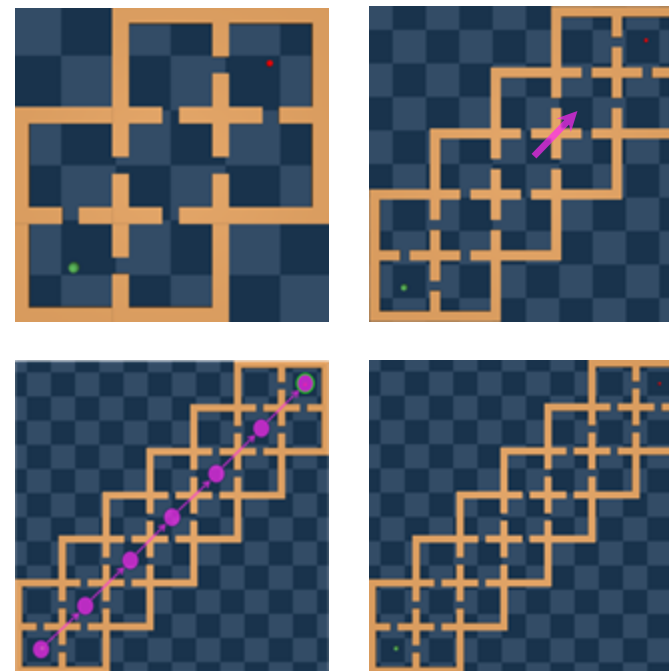
Maze4x4-Fine



Maze4x4-Coarse



Diagonal Mazes



On single-path ILGs, significantly outperforms prior approaches.
On multi-path ILGs, no other methods work stably, ILG-Learn gets 80-100%.

Guidelines on how to choose ILG task specifications for a new task

Summary and Future Work

- Summary:
 - A new, versatile LTL-style task specification for visuomotor learning of long-horizon tasks
 - Interleaved planning and RL at training time to select optimal task strategy + acquire appropriate low-level controllers
- Towards NSF SafeAI:
 - Generating ILGs from natural language descriptions using VLMs a la Lang2LTL
 - Generating safety constraints, natural for temporal logic-style ILG spec
 - Building into a system with online safety interventions and stress testing.
 - Real robots!