Planning Goals for Exploration

Edward S. Hu, Richard Chang, Oleh Rybkin, Dinesh Jayaraman

Unsupervised Task Discovery

Dropped into an unknown environment, what should an agent do to discover what it can do and how to do it?
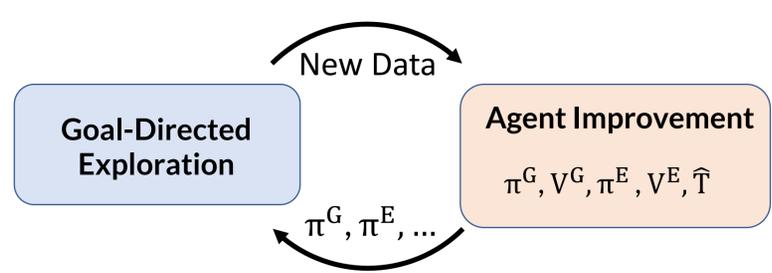

Agents first explore the environment, and then are asked to perform arbitrary tasks at test time.

More diverse exploration experience ⇒ more *generalist agents*

Task Discovery in Goal-Conditioned RL

Goal-conditioned RL: Train an agent $\pi^G(a|s,g)$ to reach any specified goal.
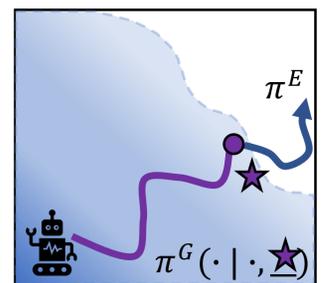
GCRL Training Loop

New Data

Goal-Directed Exploration ⟷ **Agent Improvement** $\pi^G, V^G, \pi^E, V^E, \widehat{T}$

$\pi^G, \pi^E, \ldots$

Goal Directed Exploration (Go-Explore style)

Go-phase:

Select a goal, execute π^G, and end up in s_T

Explore-phase:

Launch π^E from s_T to explore.

π^E

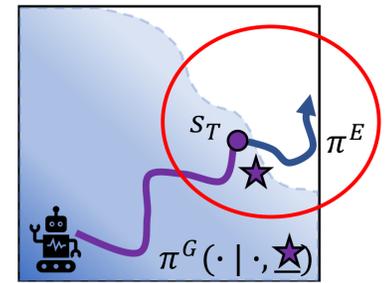$\pi^G(\cdot \mid \cdot, \bigstar)$

Open Question:

How to set goals for exploration?

Goals that maximize exploration

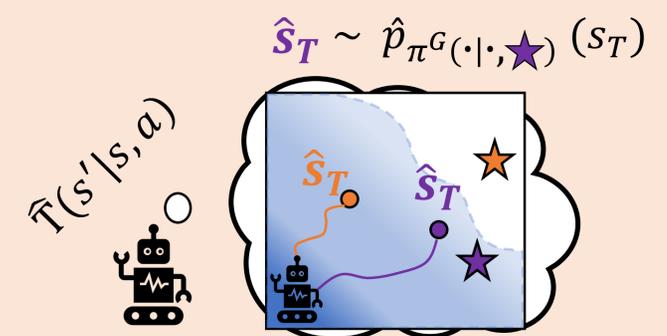Optimize for the exploration rewards gathered by our exploration policy π^E?

$$\max_g \underbrace{\mathbb{E}_{p_{\pi^G}(\cdot|\cdot,g)(s_T)}}_{\text{Where } \pi^G \text{ ends up following } g} [\underbrace{V^E(s_T)}_{\text{How much will } \pi^E \text{ explore}}]$$
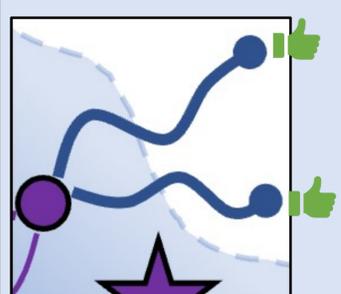
Choose goals that when followed by π^G, result in good launching points for π^E

Planning Exploratory Goals (PEG)

1. Forecast $\pi^G(\cdot \mid \cdot, g)$ trajectories

$$\hat{s}_T \sim \hat{p}_{\pi^G(\cdot|\cdot,\bigstar)}(s_T)$$

$\widehat{T}(s'|s,a)$

2. Score goals by **expl. value of terminal state**

$Score(\bigstar)$
$= V^E(\hat{s}_T)$
$= \mathbb{E}_{\pi^E}[\sum r_t^e]$
$= (👍 + 👍)/2$
$= 👍$

3. Use best goal for Go-Explore exploration

Experiments

Point Maze   Walker   Ant Maze   3-Block Stack
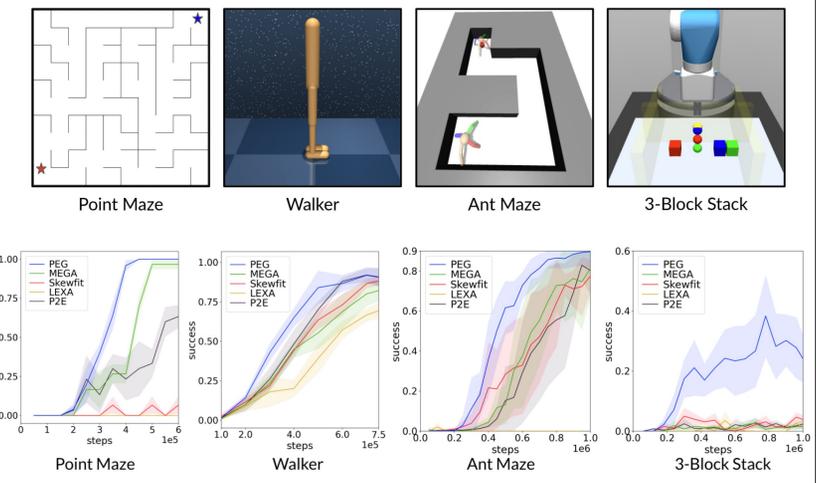
How does PEG set goals for exploration?

Point Maze   Walker

MEGA   PEG   MEGA   PEG

Ant Maze

PEG   MEGA   Skewfit   P2E
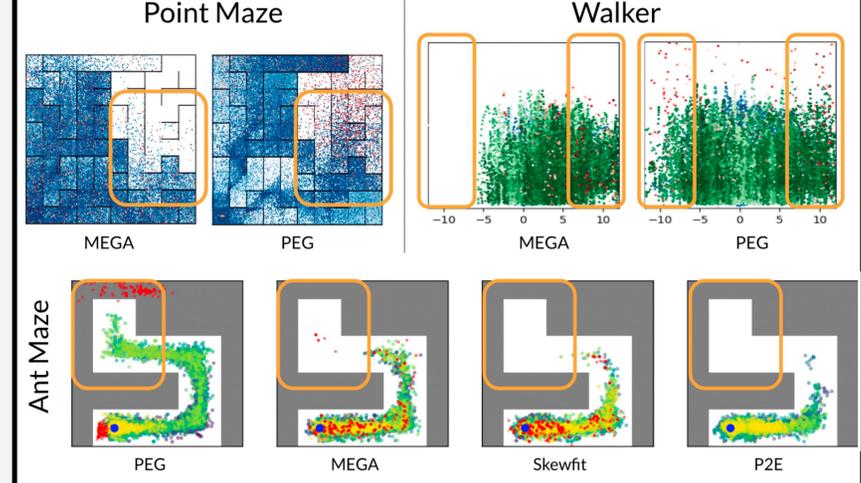
Block Stacking Exploration

Go-Phase   **Explore-Phase**
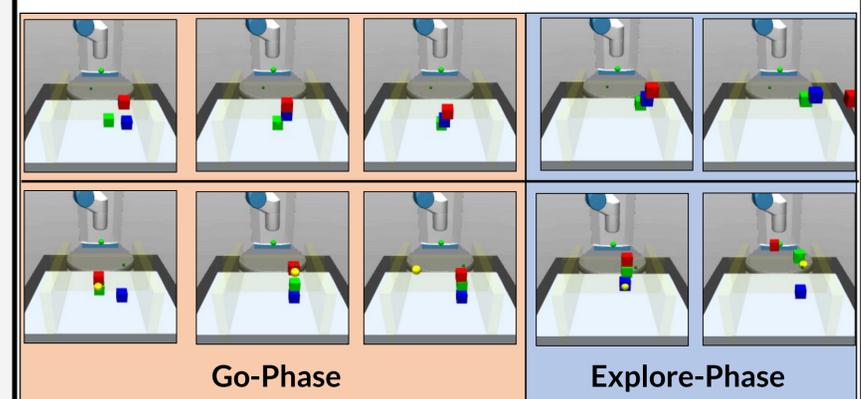
PEG commands the Go-policy π^G to stack blocks and the Explore-policy π^E throws them out.

Check out our website for more info!
https://sites.google.com/view/exploratory-goals