

Beyond Fixed Tasks: Unsupervised Environment Design for Task-Level Pairs

Daniel Furelos-Blanco¹, Charles Pert¹, Frederik Kelbel¹, Alex F. Spies¹, Alessandra Russo¹, and Michael Dennis²

¹IMPERIAL ²Google DeepMind



Motivation & Background

Task Specification via Formal Languages

- **Precise:** finite-state machines, temporal logics.
- **Training:** Uninformed sampling, a.k.a. *domain randomization (DR)*.

⚠️ DR fails when solvable task-level pairs are **rarely** sampled ⚠️

Unsupervised Environment Design (UED)

- **Adaptive curricula:** Generate levels adapted to the agent's capabilities.
- **Buffer-based (PLR, ACCEL):** Prioritize levels w/ a *score*.

⚠️ Only generates curricula over levels for a **fixed task** ⚠️

Goal

Effectively train RL agents to follow diverse instructions (**tasks**) in varied environments (**levels**)

ATLAS: Aligning Tasks and Levels for Autocurricula of Specifications

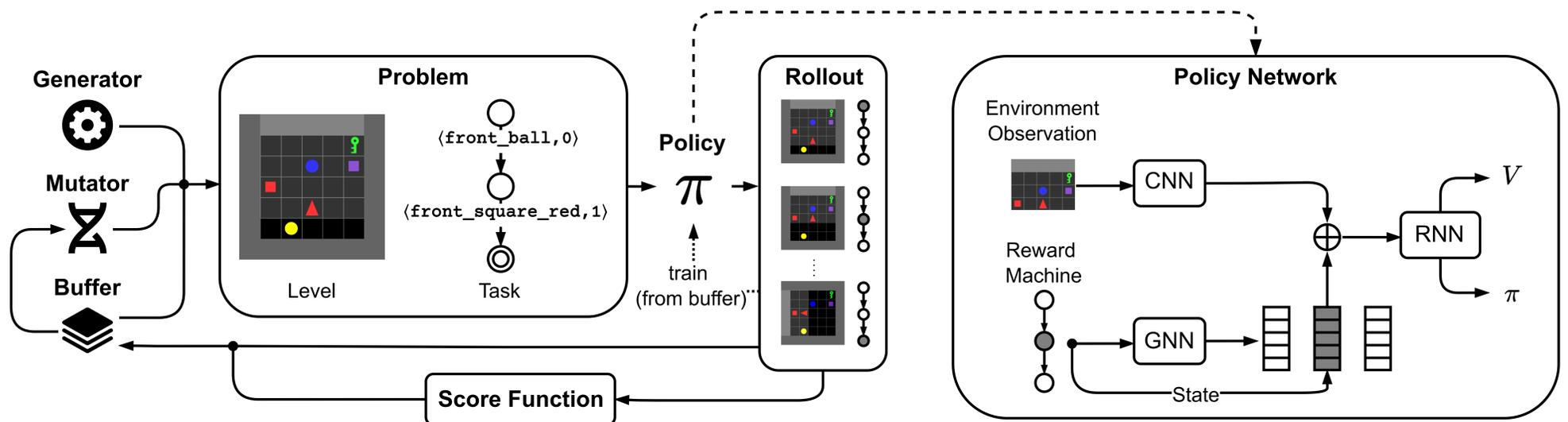
ATLAS extends UED (PLR, ACCEL) to generate curricula over **both** levels and reward machine (RM) tasks.

Joint Autocurricula

Co-designs *solvable yet challenging* RM tasks and levels

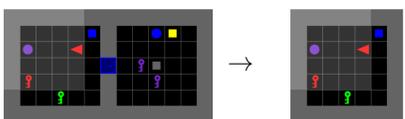
Structure-Aware Task Mutations (w/ACCEL)

Leverage *RM structure* to derive new tasks



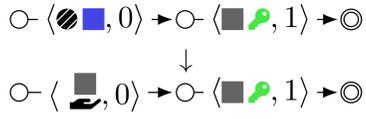
Level Edits

Add/remove rooms, move/replace objects



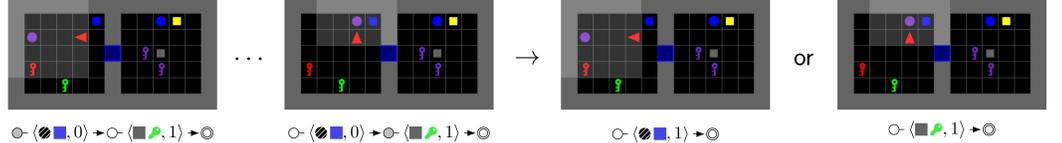
Task Edits

Add/remove states, switch proposition



Hindsight Edits

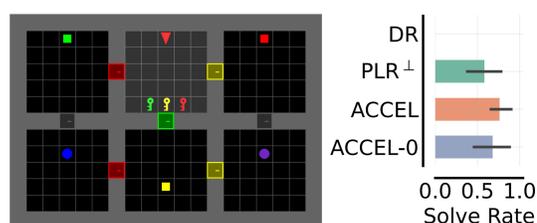
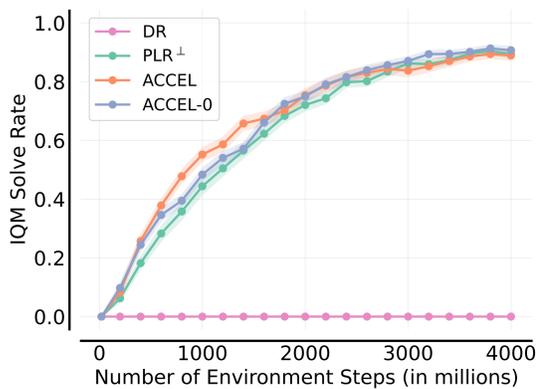
Extract preceding or succeeding problem from a rollout



Key Results

Test Performance

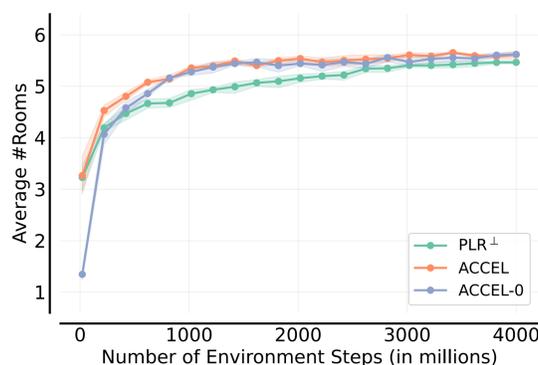
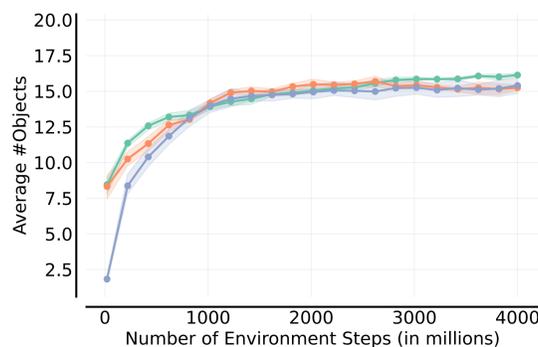
- 150 hand-designed challenging problems.
- ATLAS achieves ~90% solve rate.
- DR fails: only 2.7% solvable training pairs.



Example problem: exploration + long-term planning

Curriculum Emergence

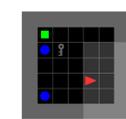
- Curricula emerge over both tasks and levels.
- ACCEL-0: curricula from simple problems via mutations alone.



Generated Training Problems

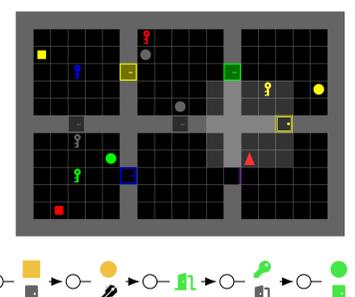
DR

Mostly unsolvable



ACCEL-0

Solvable + Challenging



Takeaways

- Joint task-level autocurricula via UED.
- Structure-aware RM mutations.
- Outperforms DR in MiniGrid-BabyAI.
- **Future:**
 - richer domains.
 - alternative scores.
 - hierarchies of RMs.

