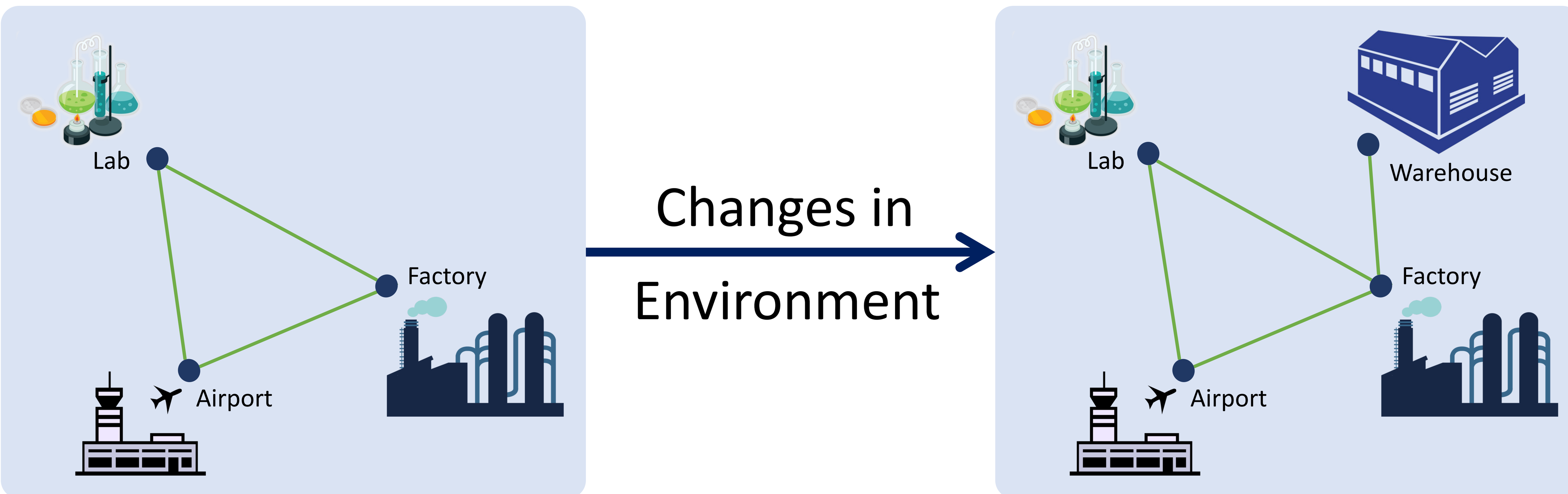
Epistemic Exploration for Generalizable Planning and Learning in Non-Stationary Stochastic Settings

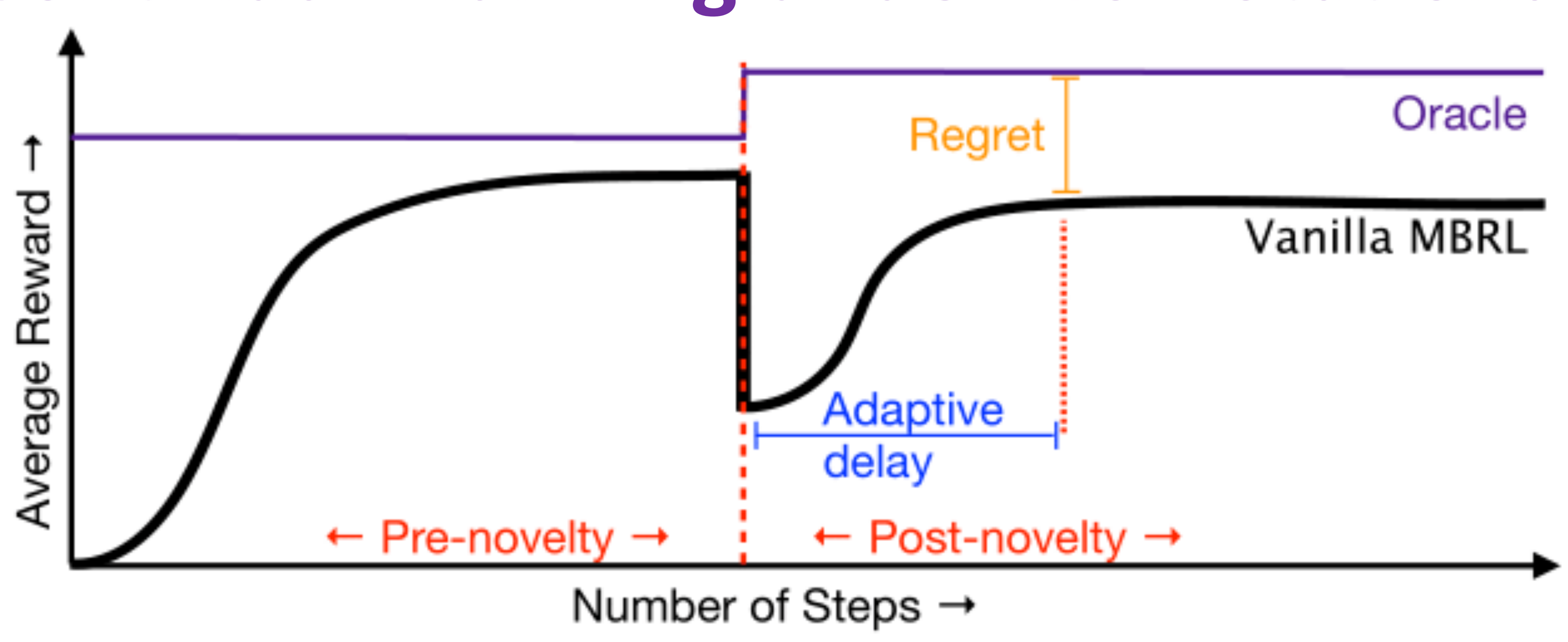
Rushang Karia*, Pulkit Verma*, Gaurav Vipat, Siddharth Srivastava

Arizona State University, AZ, USA

Motivation



Continual Planning under Non-Stationarity



Source: J. Balloch et al., NovGrid: A Flexible Grid World for Evaluating Agent Response to Novelty, AAAI Spring Symposium 2022 on Designing AI for Open Worlds

Limitations of RL in Non-Stationary Settings

- Is not sample-efficient
- Poor performance in symbolic problems

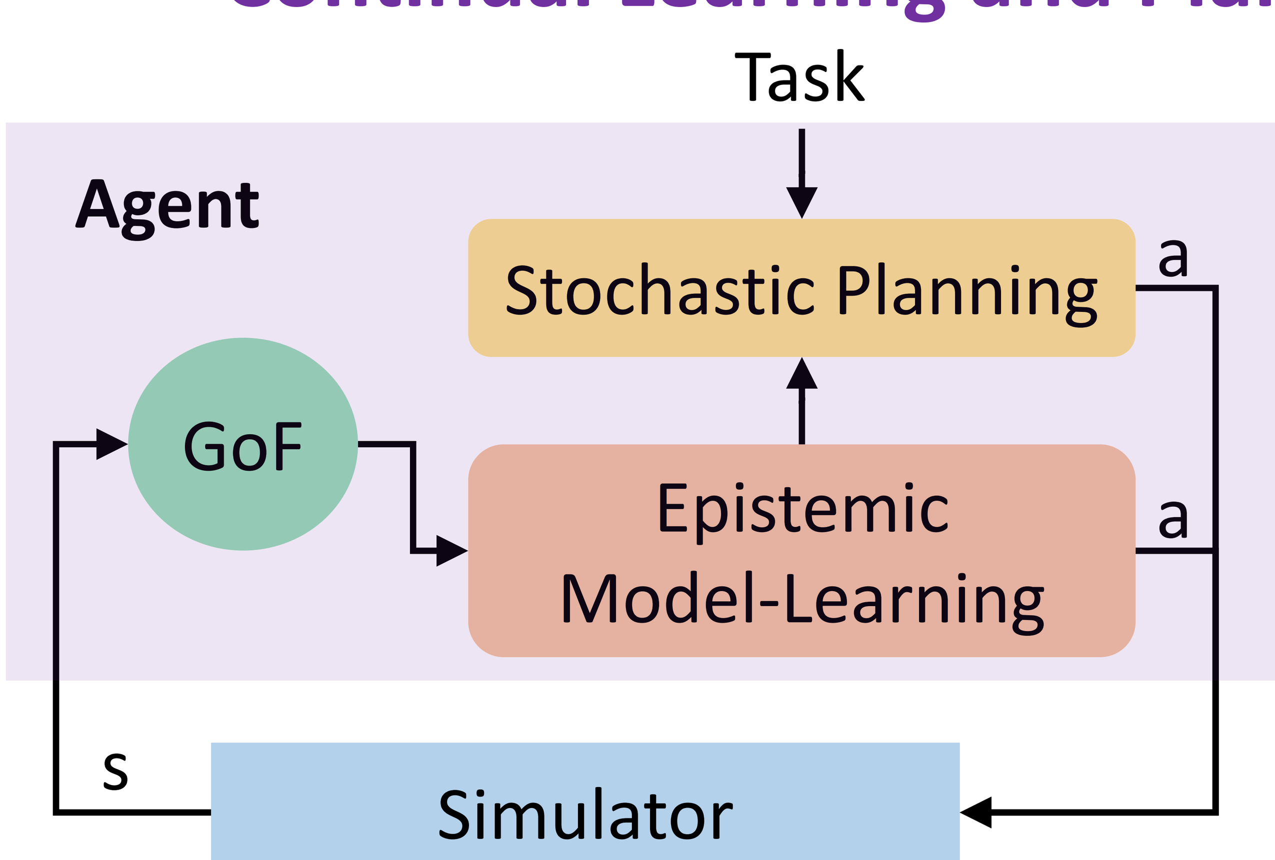
Can we learn models fast enough and use them for transfer in a sample-efficient fashion?

Types of Changes

1. Goals
2. Action effects/preconditions
3. Probability distributions

Can we reduce Adaptive Delay and Regret across all types of changes?

Continual Learning and Planning (CLaP)



Model Learning
Active
Query-based
Autonomous
Capability
Estimation*

Key Features of CLaP

$$V^*(s) = \max_a \left[R(s,a) + \gamma \sum_{s'} \delta(s,a,s')\, V^*(s') \right]$$

- ✓ Learns **models** for stochastic planning
- ✓ Performs **investigative exploration** to resolve discrepancies in current model
- ✓ **Goodness-of-fit** tests for o.o.d. changes
- ✓ **Theoretical guarantees** of convergence

*P. Verma, R. Karia, S. Srivastava, Autonomous Capability Assessment of Sequential Decision-Making Systems in Stochastic Settings. NeurIPS 2023.

Results

We found that CLaP results in **significantly better**
(a) Sample Efficiency (b) Average Reward (c) Adaptive delay