A Domain-Independent Agent Architecture for Adaptive Operation in Evolving Open Worlds

Shiwali Mohan [1], Wiktor Piotrowski [1], Roni Stern [2], Sachin Grover [1], Sookyung Kim [1], Jacob Le [1], Johan de Kleer [1], Yoni Sher [1]

[1] Stanford Research Institute
[2] Ben Gurion University of the Negev

Abstract Reprint. This is an abstract reprint of a journal article by Mohan, Piotrowski, Stern, Grover, Kim, Le, Sher, and de Kleer (2024).

Abstract

Model-based reasoning agents are ill-equipped to act in novel situations in which their model of the environment no longer sufficiently represents the world. We propose HYDRA, a framework for designing model-based agents operating in mixed discrete-continuous worlds that can autonomously detect when the environment has evolved from its canonical setup, understand how it has evolved, and adapt the agents' models to perform effectively. HYDRA is based upon PDDL+, a rich modeling language for planning in mixed, discrete-continuous environments. It augments the planning module with visual reasoning, task selection, and action execution modules for closed-loop interaction with complex environments. HYDRA implements a novel meta-reasoning process that enables the agent to monitor its own behavior from a variety of aspects. The process employs a diverse set of computational methods to maintain expectations about the agent's own behavior in an environment. Divergences from those expectations are useful in detecting when the environment has evolved and identifying opportunities to adapt the underlying models. HYDRA builds upon ideas from diagnosis and repair and uses a heuristics-guided search over model changes such that they become competent in novel conditions. The HYDRA framework has been used to implement novelty-aware agents for three diverse domains - CartPole++ (a higher dimension variant of a classic control problem), Science Birds (an IJCAI competition problem), and PogoStick (a specific problem domain in Minecraft). We report empirical observations from these domains to demonstrate the efficacy of various components in the novelty meta-reasoning process.

References

Mohan, S.; Piotrowski, W.; Stern, R.; Grover, S.; Kim, S.; Le, J.; Sher, Y.; and de Kleer, J. 2024. A domain-independent agent architecture for adaptive operation in evolving open worlds. *Artificial Intelligence*, 334: 104161.