The Thirty-Eighth AAAI Conference on Artificial Intelligence (AAAI-24)

Learning Generalizable and Composable Abstractions for Transfer in Reinforcement Learning

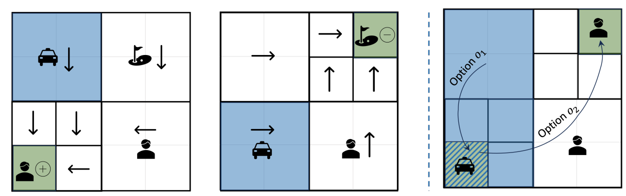
Rashmeet Kaur Nayyar
School of Computing and Augmented Intelligence
Arizona State University, Tempe, AZ, USA, 85281
rmnayyar@asu.edu

Abstract

Reinforcement Learning (RL) in complex environments presents many challenges: agents require learning concise representations of both environments and behaviors for efficient reasoning and generalizing experiences to new, unseen situations. However, RL approaches can be sample-inefficient and difficult to scale, especially in long-horizon sparse reward settings. To address these issues, the goal of my doctoral research is to develop methods that automatically construct semantically meaningful state and temporal abstractions for efficient transfer and generalization. In my work, I develop hierarchical approaches for learning transferable, generalizable knowledge in the form of symbolically represented options, as well as for integrating search techniques with RL to solve new problems by efficiently composing the learned options. Empirical results show that the resulting approaches effectively learn and transfer knowledge, achieving superior sample efficiency compared to SOTA methods while also enhancing interpretability.

1 Introduction

The focus of my doctoral research is on learning abstractions for effective transfer and generalization in planning and reinforcement learning. The guiding question of my research is: *"How do intelligent agents learn succinct and meaningful abstract knowledge of the world to solve related new problems efficiently?"*. Consider most self-driving cars of today that train and over-fit to known scenarios which makes it difficult to start operations in new cities and countries. Despite many efforts in autonomously learning state (Whiteson 2010) and temporal (Klissarov and Precup 2021) abstractions, these approaches suffer from limited reusability and sample-efficiency. My dissertation aims to create general hierarchical algorithms that learn reusable abstractions and employ abstractions that are automatically identified as useful for solving new problems. To realize this goal, I propose learning and using combined state and temporal abstractions or options with symbolic representations that satisfy the desirable properties: interpretability, composability, and generalizability (as illustrated in Fig. 1).



(a) Option o_1 (b) Option o_2 (c)

Figure 1: Illustration of options in taxi world with passengers and a destination. The images (a) and (b) show abstract policies and initiation sets for option o_1 (to navigate and pickup passenger) and option o_2 (to navigate and dropoff passenger) that are generalizable due to abstraction. The image (c) shows composability of the options as abstract state in the termination of o_1 (shown in green) are in the initiation of o_2 (shown in blue).

2 Proposed Approach

My dissertation endeavors to learn abstractions that satisfy the desirable properties with a focus on algorithms for improving sample-efficiency and transfer.

2.1 Learning State Abstractions

Our recently published work (Dadvar, Nayyar, and Srivastava 2023) introduces a novel top-down method called CAT+RL to learn a hierarchy of state abstractions in the form of a Conditional Abstraction Tree (CAT) while carrying out RL. The CAT+RL algorithm takes as input a Stochastic Shortest Path problem (SSP) and effectively learns a CAT and an abstract policy. My follow-up work further aims to build upon this result to address the problem of learning symbolically represented temporal abstractions or options to improve transfer and generalization to new unseen problems as described in Sec. 2.2.

2.2 Learning Temporal Abstractions

Our recent research presents a novel method called **C**AT **O**ptions **Plan**ning and **Learn**ing -- COPlanLearn -- for Transfer Reinforcement Learning (RL) for SSP problems in factored domains (Nayyar, Verma, and Srivastava 2023). It learns a library of options with abstract representations and recomposes these options while discovering novel options by integrating planning with RL to solve new problems.

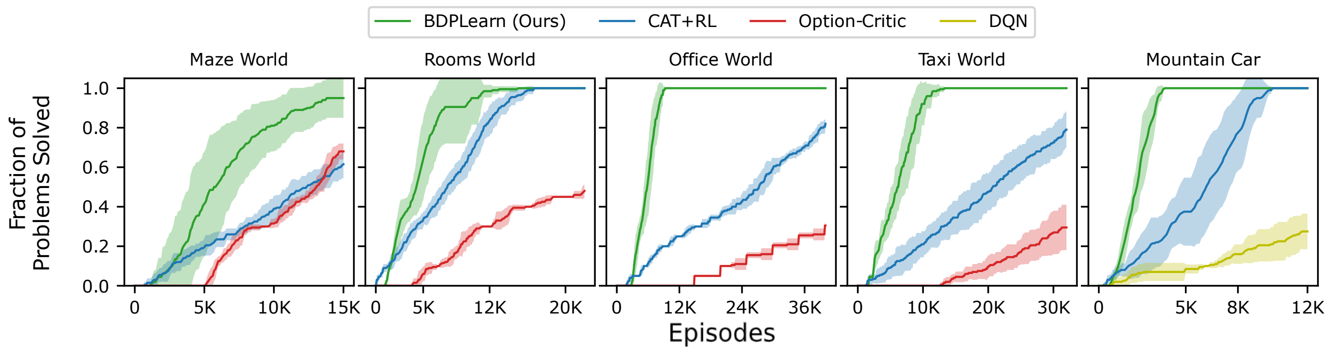


Figure 2: An average of the fraction of problems solved vs learning episodes required by each approach in each domain, computed using 10 independent trials. A total of 20 problems (2 source, 18 target) were solved sequentially by each approach.

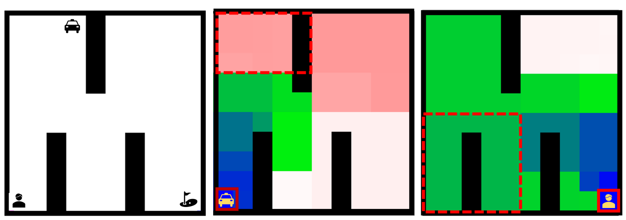


Figure 3: Visualization of options for Taxi World. Dashed lines show the initiation set and solid lines show the termination set for an option. The left image shows the initial state for the taxi and the passenger, and the destination. The middle image shows an option to navigate and pickup the passenger. The right image shows an option to navigate and dropoff the passenger.

The presented approach first learns a semantically well-defined state abstraction from an input problem instance and then uses this abstraction to invent high-level options, to learn abstract policies for executing them, as well as to create abstract symbolic representations for representing them. Given a new problem instance, our overall approach conducts a novel bi-directional search over the learned option representations while also inventing new options as needed. Our main contributions are approaches for learning transferable, generalizable knowledge in the form of symbolically represented options, as well as for integrating search techniques with RL to solve new problems by efficiently composing the learned options.

3 Preliminary Results

Extensive empirical evaluations on CAT+RL (Dadvar, Nayyar, and Srivastava 2023) demonstrated that it effectively draws out similarities across the state space and enables the vanilla Q-learning algorithm to outperform SOTA methods. Furthermore, empirical results on COPlanLearn (Fig. 2) show that the resulting approach effectively transfers learned knowledge to problems with different initial and goal configurations, achieving superior sample efficiency compared to SOTA methods. A key advantage over Option-Critic (Bacon, Harb, and Precup 2017) is that it does not need to predefine the number of options. The approach also effectively transfers from less to more cluttered and from smaller to larger environments. Visualizations of learned options in the Taxi World (Fig. 3) exhibit desirable properties of interpretability, composability, and generalizability.

4 Future Directions

With the right kind of additional analysis, I aim to extend the proposed methods to invent options with an increased scope of their applicability. Specifically, I plan to investigate learning relational options that are transferable to diverse settings and are generalizable to an increasing number of objects in RL, as previously achieved in planning (Karia, Nayyar, and Srivastava 2022). Moreover, my work so far focuses on domains with discrete action spaces. I am interested in extending our work to the case where the underlying action space is continuous. These ideas can be extended to learn abstract high-level action models and identify drift in them for related environments, similar to (Nayyar, Verma, and Srivastava 2022). Lastly, my goal is to investigate the connections between hierarchies of state and temporal abstractions and support the resulting algorithms with theoretical guarantees.

References

Bacon, P.-L.; Harb, J.; and Precup, D. 2017. The Option-Critic Architecture. In *Proc. AAAI*.

Dadvar, M.; Nayyar, R. K.; and Srivastava, S. 2023. Conditional Abstraction Trees for Sample-efficient Reinforcement Learning. In *Proc. UAI*.

Karia, R.; Nayyar, R. K.; and Srivastava, S. 2022. Learning Generalized Policy Automata for Relational Stochastic Shortest Path Problems. In *Proc. NeurIPS*.

Klissarov, M.; and Precup, D. 2021. Flexible Option Learning. In *Proc. NeurIPS*.

Nayyar, R. K.; Verma, P.; and Srivastava, S. 2022. Differential Assessment of Black-box AI Agents. In *Proc. AAAI*.

Nayyar, R. K.; Verma, S.; and Srivastava, S. 2023. Learning Generalizable Symbolic Options for Transfer in Reinforcement Learning. In *NeurIPS GenPlan Workshop*.

Whiteson, S. 2010. Adaptive Tile Coding. *Adaptive Representations for Reinforcement Learning*, 65–76.