Dynamic probabilistic logic models for effective task-specific abstractions in RL

Harsha Kokel
The University of Texas at Dallas
hkokel@utdallas.edu

Sriraam Natarajan
The University of Texas at Dallas
Sriraam.Natarajan@utdallas.edu

Balaraman Ravindran
Robert Bosch Centre for
Data Science and Artificial Intelligence
Indian Institute of Technology Madras
ravi@cse.iitm.ac.in

Prasad Tadepalli
Oregon State University
tadepall@eecs.oregonstate.edu

Abstract

In many real-world domains, e.g., driving, the state space of offline planning is rather different from the state space of online execution. Planning typically occurs at the level of deciding the route, while online execution needs to take into account dynamic conditions such as locations of other cars and traffic lights. The agent typically does not have access to the dynamic part of the state at the planning time, e.g., future locations of other cars, nor does it have the computational resources to plan an optimal policy in advance that works for all possible traffic events. The key principle that enables agents to deal with these informational and computational challenges is abstraction. In the driving example, the high level state space consists of coarse locations such as "O'hare airport" and high level actions such as take "Exit 205," while the lower level state space consists of a more precise location and velocity of the car and actions such as turning the steering wheel. Importantly, excepting occasional unforeseen failures, the two levels operate independently of each other and depend on different kinds of information available at different times.

To achieve this, we investigate the integration of planning and RL in a hierarchical framework called RePReL. We adapt dynamic probabilistic logic model to specify bisimilarity conditions of the MDPs to obtain safe and effective task-specific state abstractions. Our empirical evaluations, in a grid world domain and a robotic task, show that such abstractions can result in efficient learning and effective transfer.

Keywords: Reinforcement Learning, Planning, Hierarchical RL, Abstraction, Statistical Relational AI

Acknowledgements

This work is supported by ARO award W911NF2010224 (HK & SN), AFOSR award FA9550-18-1-0462 (SN), DARPA contract N66001-17-2-4030 (PT) and NSF grant IIS-1619433 (PT). Any opinions, findings, conclusion or recommendations expressed in this material are those of the authors and do not necessarily reflect the view of the ARO, AFOSR, NSF, DARPA or the US government.