Di Marzo Gabriele Giuseppe
ID: 2012633

Model-based lifelong RL: abstraction of transition model through several MDP

Lifelong RL address the problem of generalization of any task. Through a implementation of Bayesian framework for different MDP, the goal of the paper is find a way to link each MDP.

Each of this MDP are related to the same task family, the only difference is in the agent structure (half-cheetah, hopper, humanoid …).

The problem is building up through a Hidden-parameters MDP (HiP-MDP), and approximate 2 probability distribution:

- Pr(W) = world model posterior
- Pr(w_i) = task-specific posterior

As the learner is exposed to more and more tasks, this posteriors should converge to the real distribution.

Each time the agent encounters a new task, the task-specific model it's initialize using the world-model posterior and further train it with data collected only from the new task.

With a model that is able to capture different levels of the uncertainty within HiP-MDPs, an agent can employ sample-based Bayesian exploration to further improve sample-efficiency.

main reference:

- Model-based Lifelong Reinforcement Learning with Bayesian Exploration (2022)

others:

- Hidden Parameter Markov Decision Processes: A Semiparametric Regression Approach for Discovering Latent Task Parametrizations (2013)
- Environments for Lifelong Reinforcement Learning (2018)