Hybrid Deep RePReL
Integrating Relational Planning and Reinforcement Learning for Information Fusion
Motivation

Environment

Action

Observation

Agent

* Assumed homogeneous and obtained from a single source
Motivation

- Source of information may differ
- Type of data might be different
**Given:** A sequential decision making problem with a combination of structured and unstructured data.

**To Do:** Develop a hybrid architecture that learns to act.
Structured Data:
- Symbolic representations like tabular data, predicate logic, knowledge graph, etc

Unstructured Data:
- Raw, free-form data like text, image, audio, etc
Structured Data:

- Passenger’s details
  \[ \text{at}(p1,l1), \ \text{dest}(p1,d1) \]
  \[ \text{at}(p2,l2), \ \text{dest}(p1,d2) \]

Unstructured Data:

- Taxi Location
- Geography

from images
RePReL

Symbolic Planner

Abstraction Reasoner

Reinforcement Learners

initial state

state, reward

action

Kokel et al ICAPS 2021
RePReL

Goal directed relational MDP:
\[ <S, A, P, R, \gamma, G> \]
RePReL

- Plan the sequence of high level subgoals and learn to execute each subgoal at lower level

- Advantage:
  - Compositionality
  - Task specific state representations

- Dynamic First Order Conditional Influence statements to obtain task-specific abstract representations

Kokel et al ICAPS 2021
Hybrid Deep RePReL

D-FOCI

Symbolic Planner

Abstraction Reasoner

RL agents

Initial state

state

action

Environment
Hybrid Deep RePReL

Initial state

Symbolic Planner

subgoals

Abstraction Reasoner

Input pre-processor

Merge

RL agents

Environment

D-FOCI
HD RePReL Learning

- Initialize buffers
- Get high level plan
- For each subgoal
  - Loop till the subgoal is achieved or # steps exceeds
    - Get the abstract state
    - Get the policy for that subgoal
    - Take a step and observe reward, next state
    - Add <S, A, R, S> to the buffer
- Update the subgoal policy using samples from the buffers
Experiments

- Sample efficiency
- Generalization across objects

- transport one passenger
- Transport two passengers
- Transport three passengers
Experiments

- Sample efficiency
- Generalization across objects

- Make Bread
- Build a house
- Break a rock
Summary

- Combined a symbolic planner with a Deep RL agent for information fusion
- Provide a batch learning algorithm for RePReL framework
- Demonstrate **sample efficiency**, that is significant reduction in the number of steps required for the model to learn an optimal policy for the task
- Demonstrate **efficient generalization** over number of objects
- Provide hybrid approach for structured and unstructured data
QUESTIONS?
THANKS