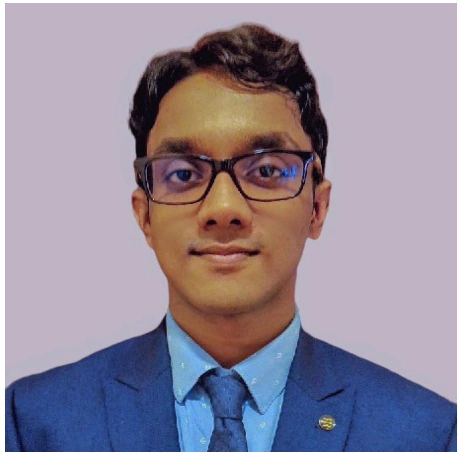
Hybrid Deep RePReL

Integrating Relational Planning and Reinforcement Learning for Information Fusion

UT Dallas
IIT Madras
AFRL — The Air Force Research Laboratory
Oregon State University

Harsha Kokel

UT DALLAS



Nikhilesh Prabhakar

UT DALLAS

Balaraman Ravindran

IIT MADRAS

Erik Blasch

AFRL THE AIR FORCE RESEARCH LABORATORY

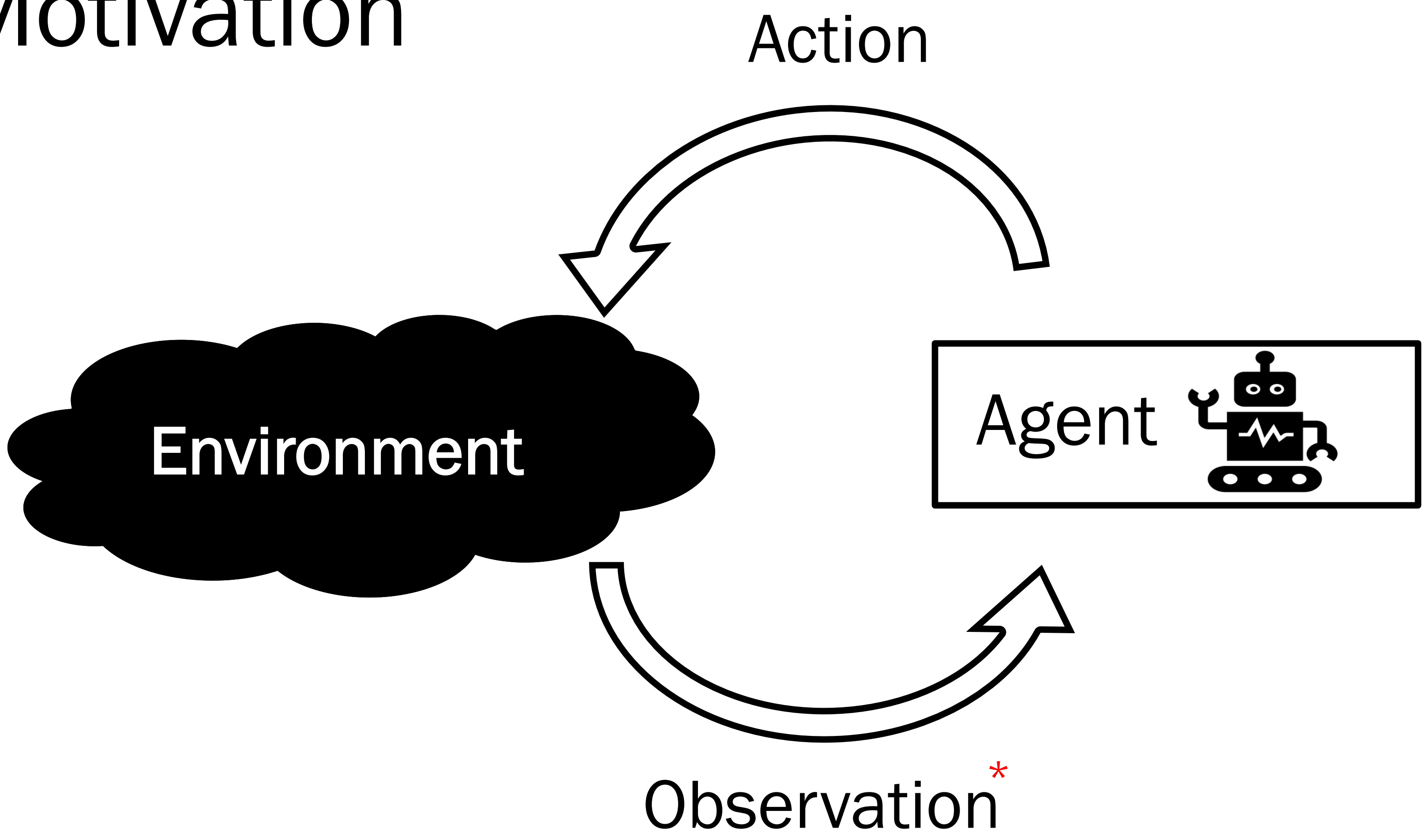
Prasad Tadepalli

Oregon State University

Sriraam Natarajan

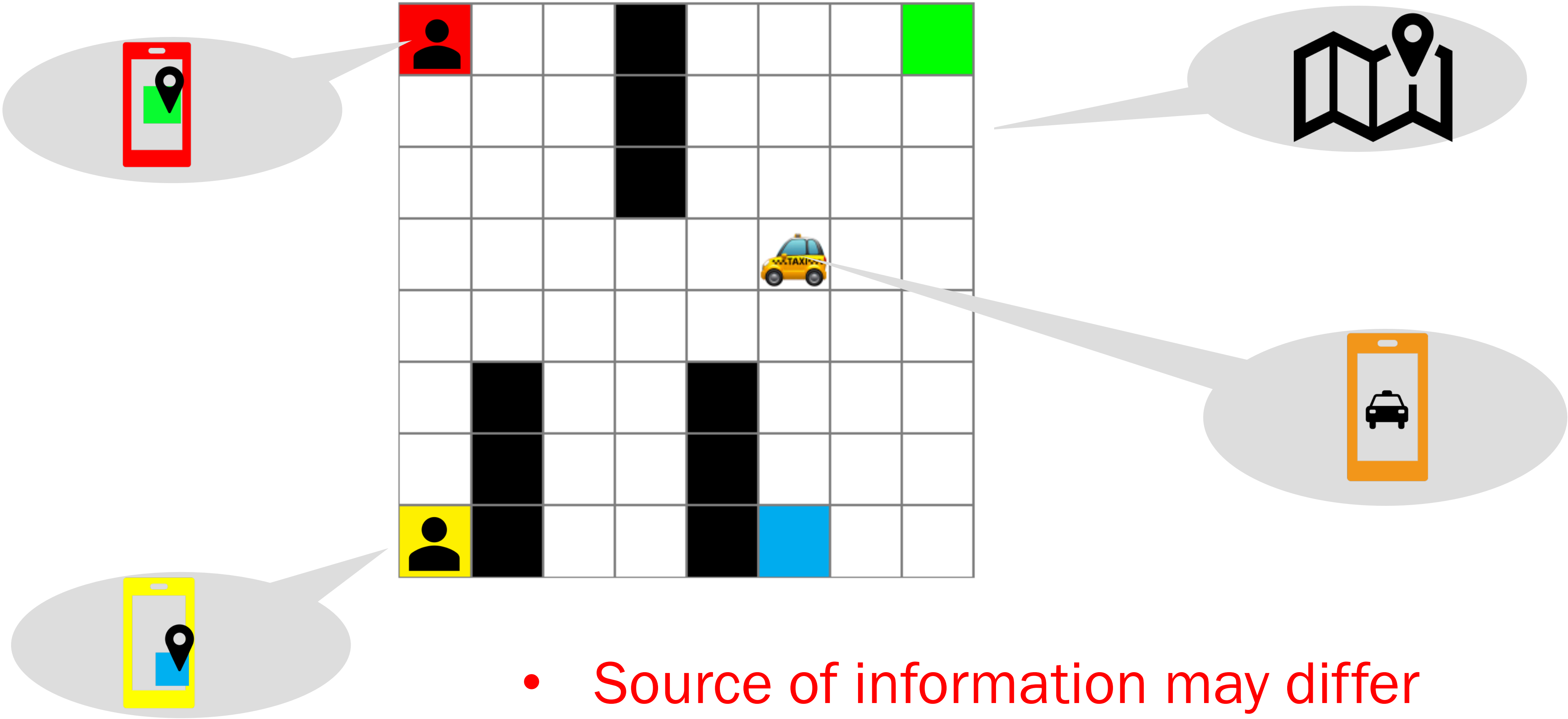
UT DALLAS

Motivation



* 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:

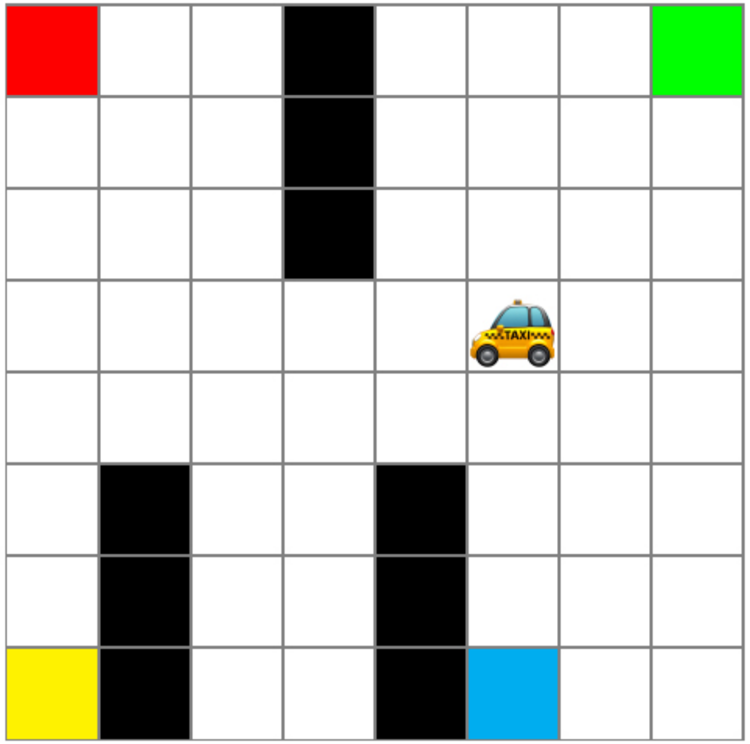
- Passenger's details

```
at(p1,l1), dest(p1,d1)
at(p2,l2), dest(p1,d2)
```

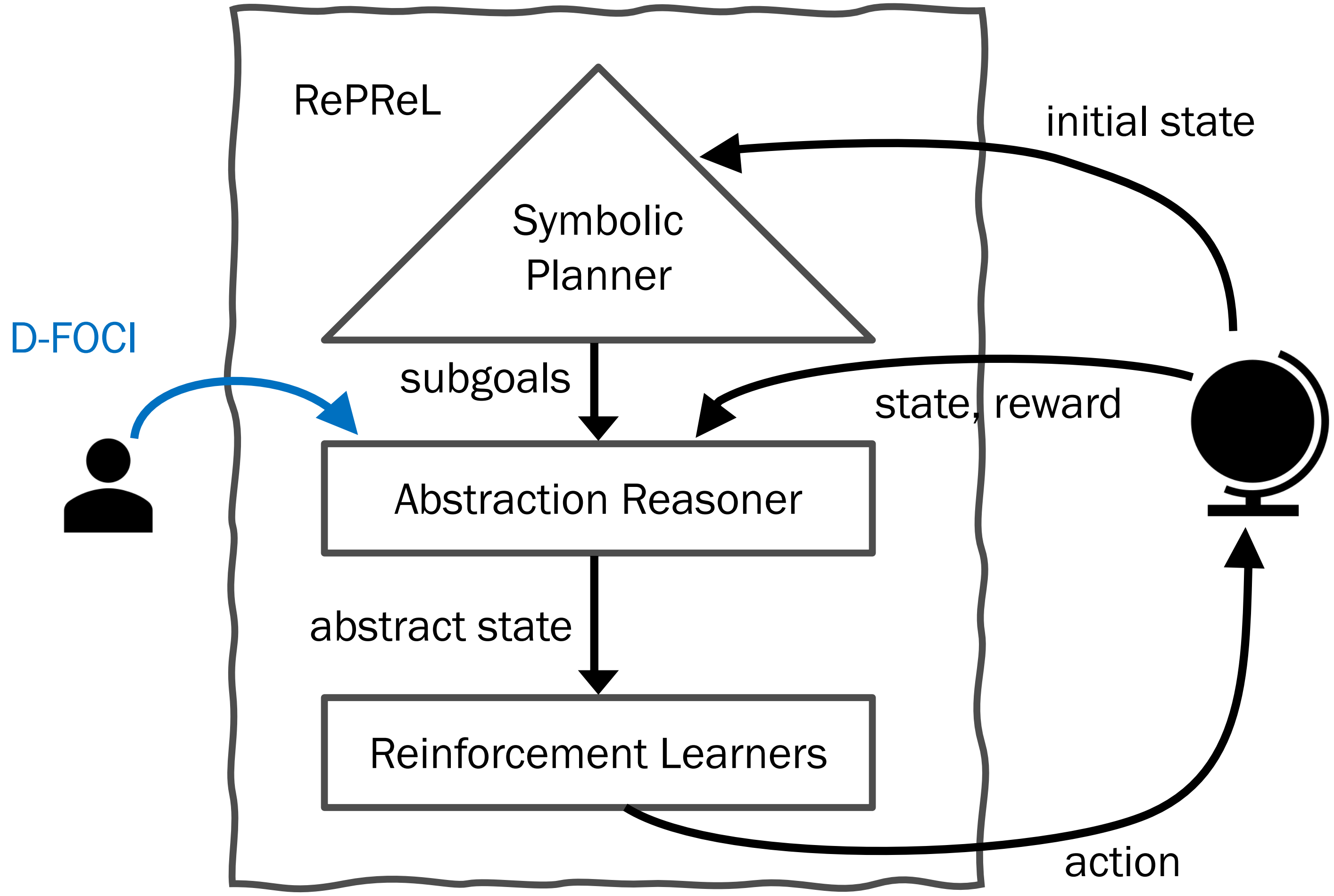
Unstructured Data:

- Taxi Location
- Geography

from images



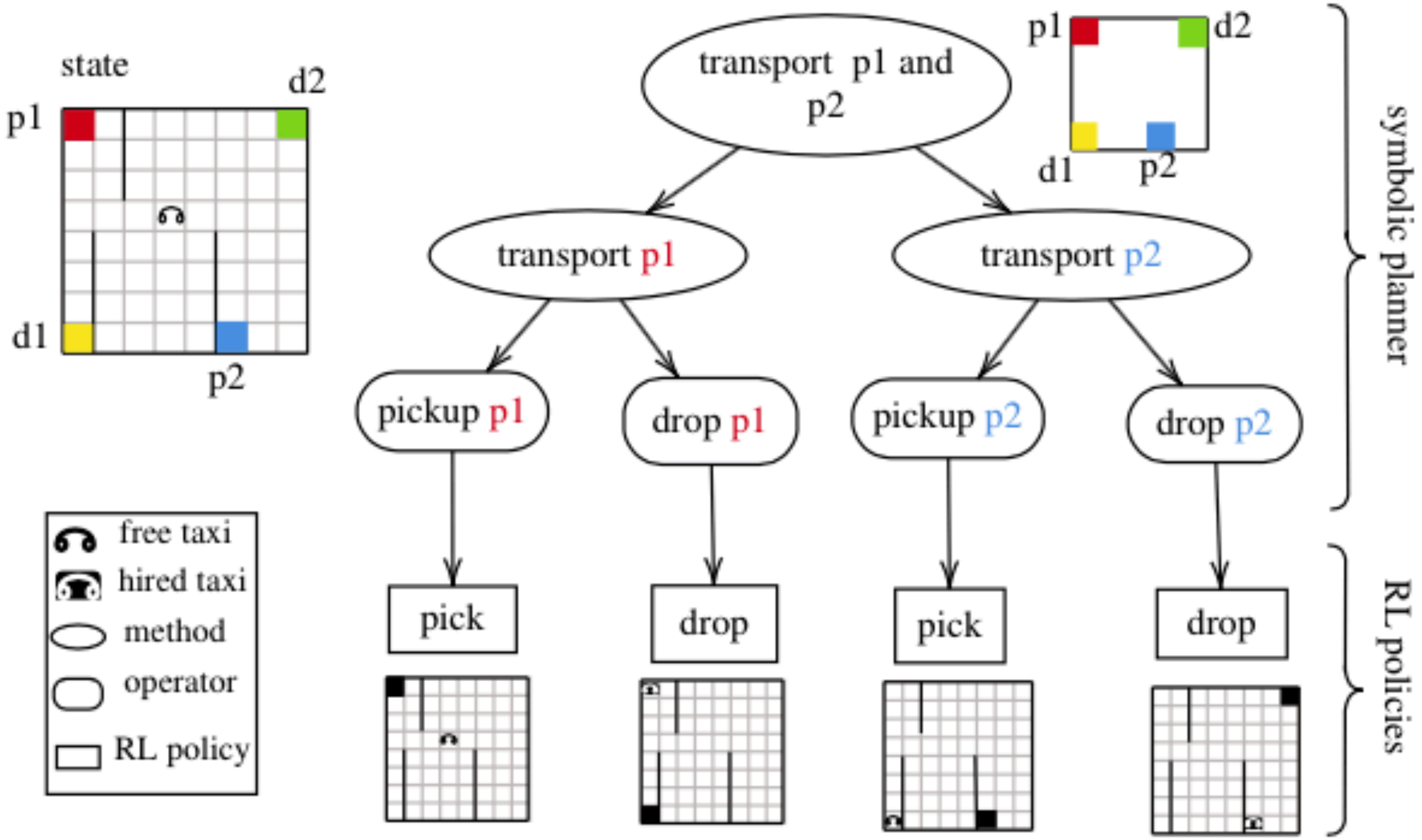
RePReL



Kokel et al ICAPS 2021

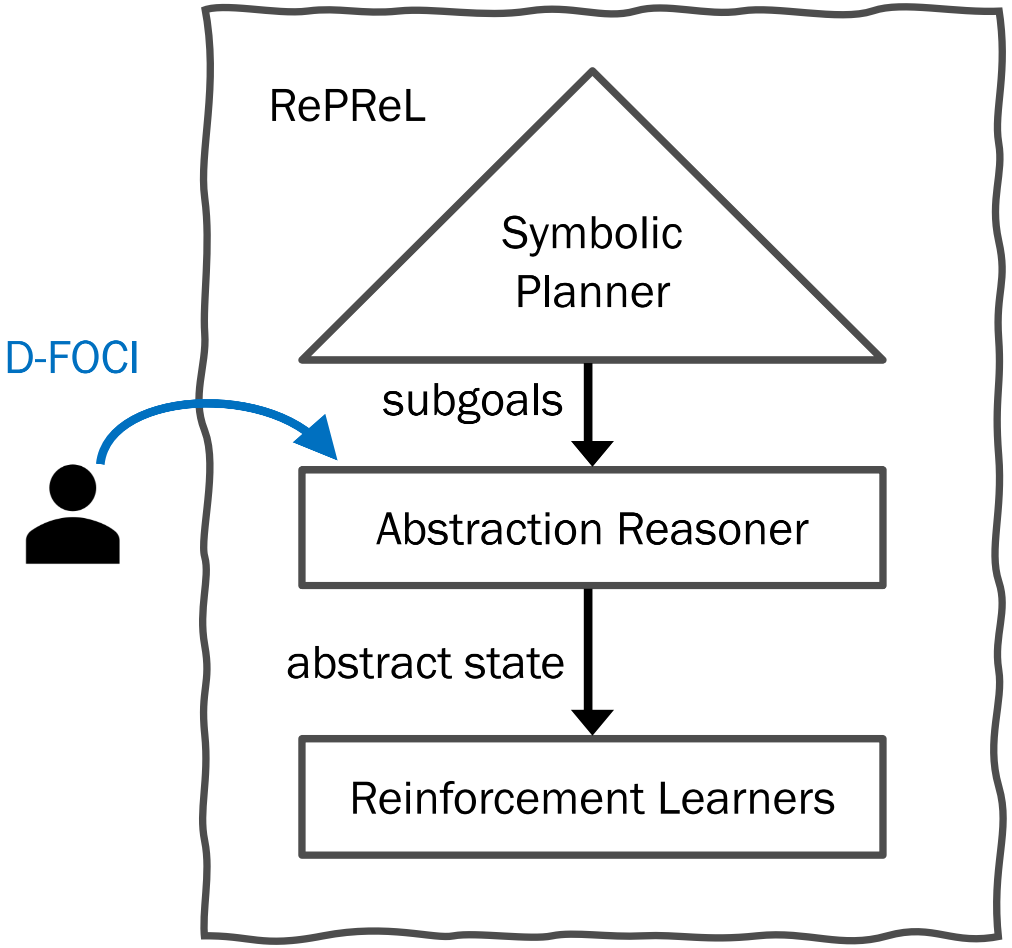
RePReL

Goal directed relational MDP:
<S, A, P, R, γ, G>



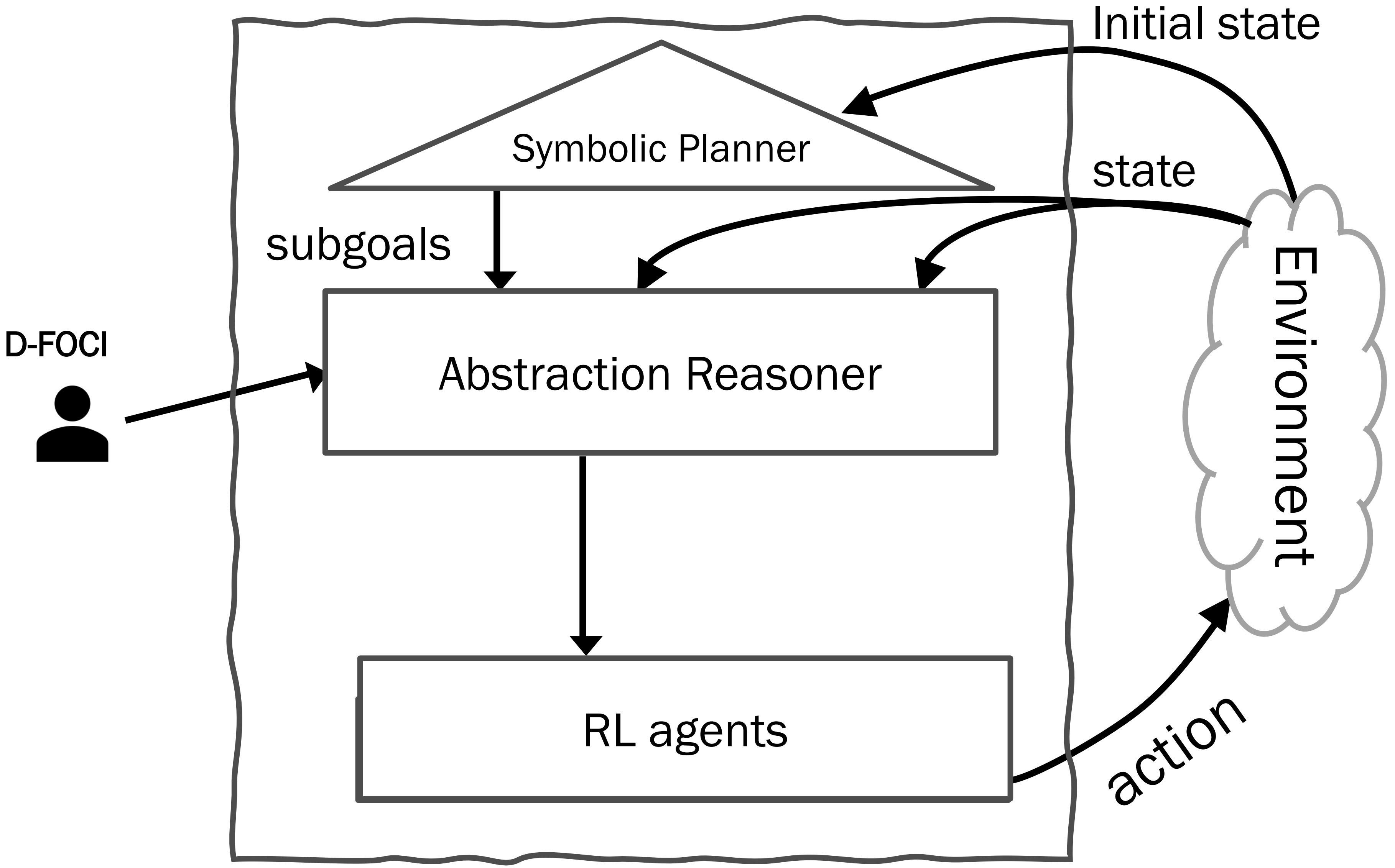
RePReL

- 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

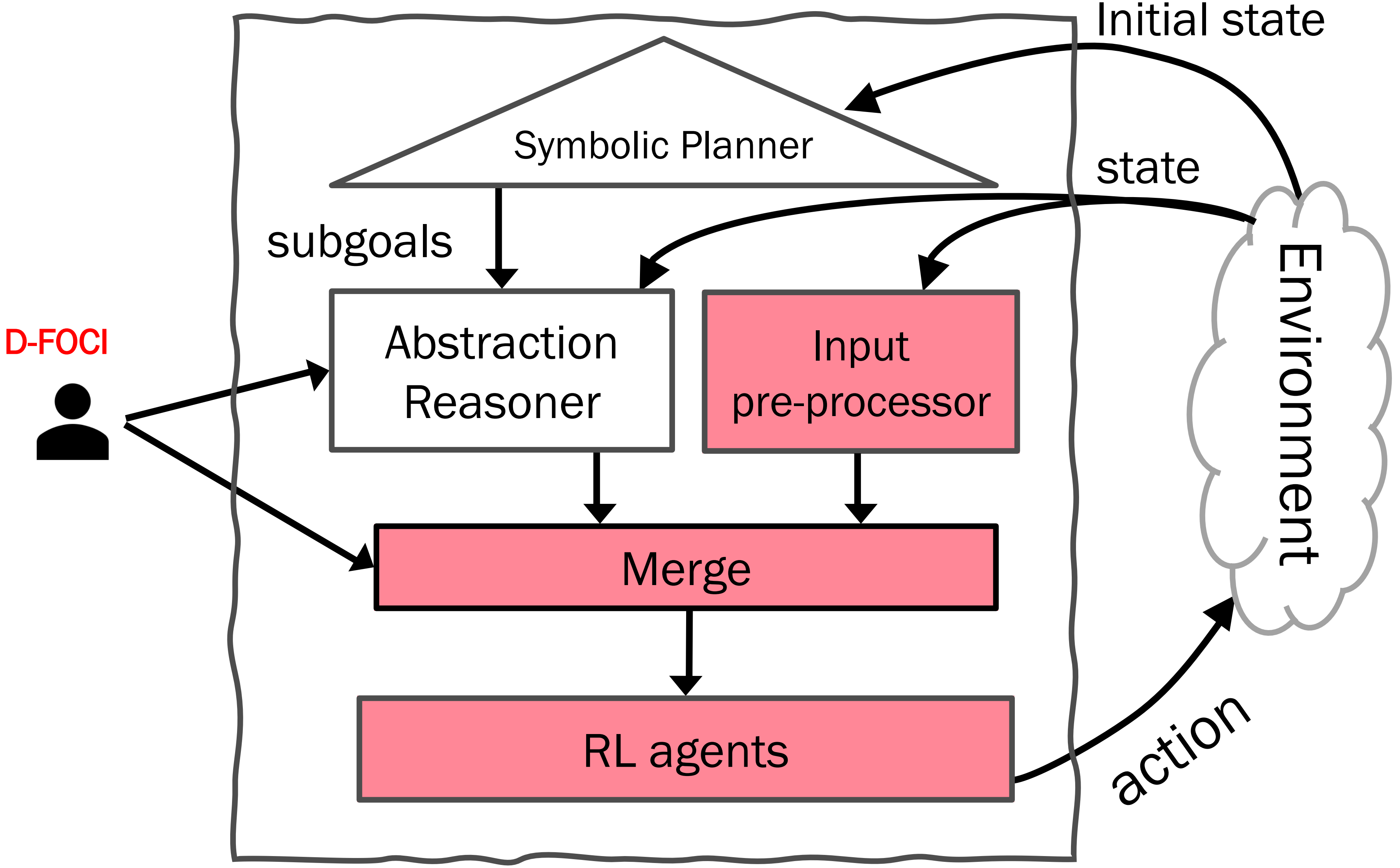


Kokel et al ICAPS 2021

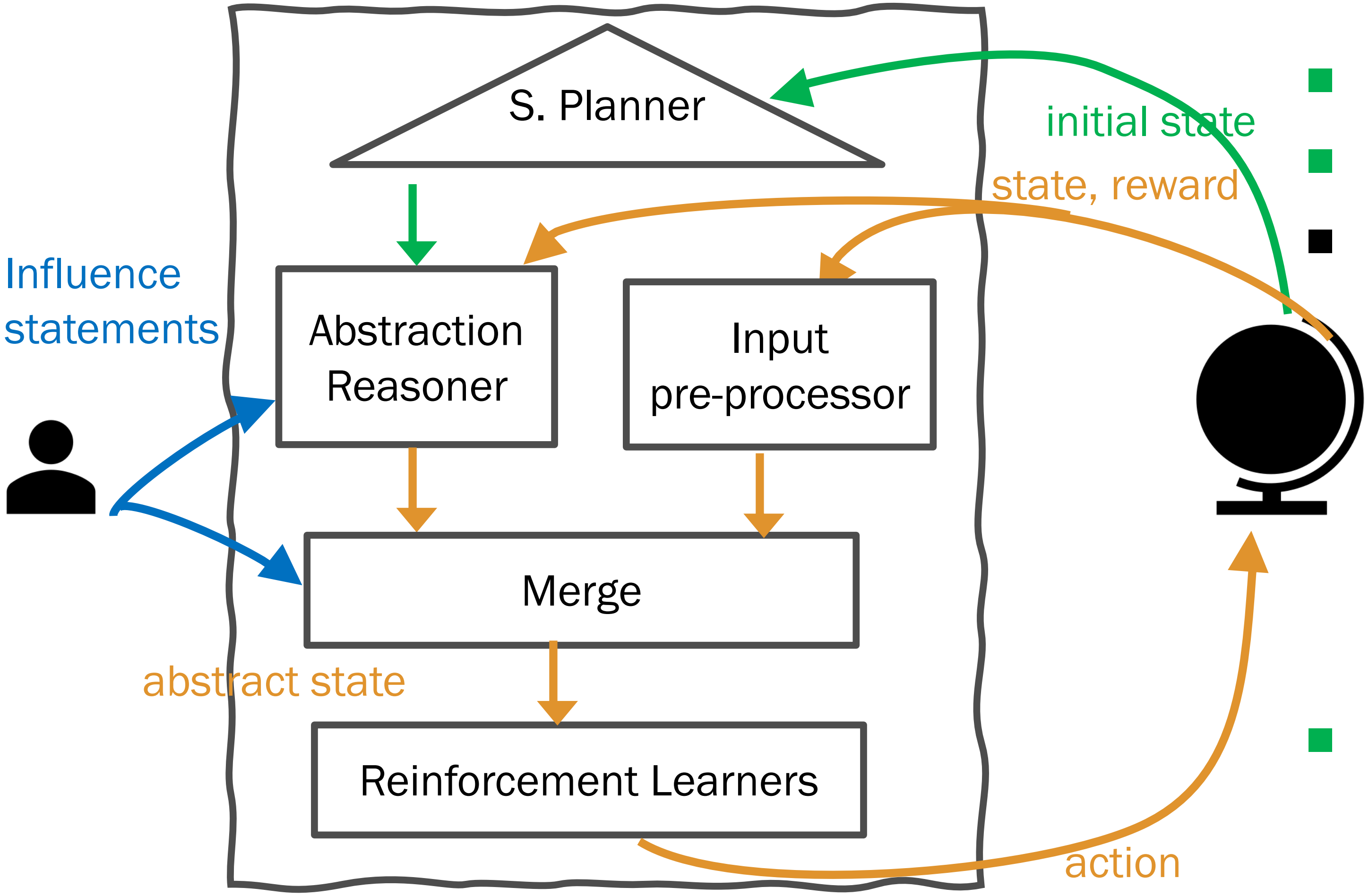
Hybrid Deep RePReL



Hybrid Deep RePReL



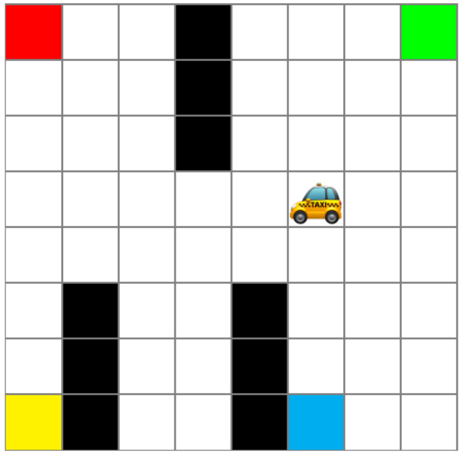
HD RePReL Learning

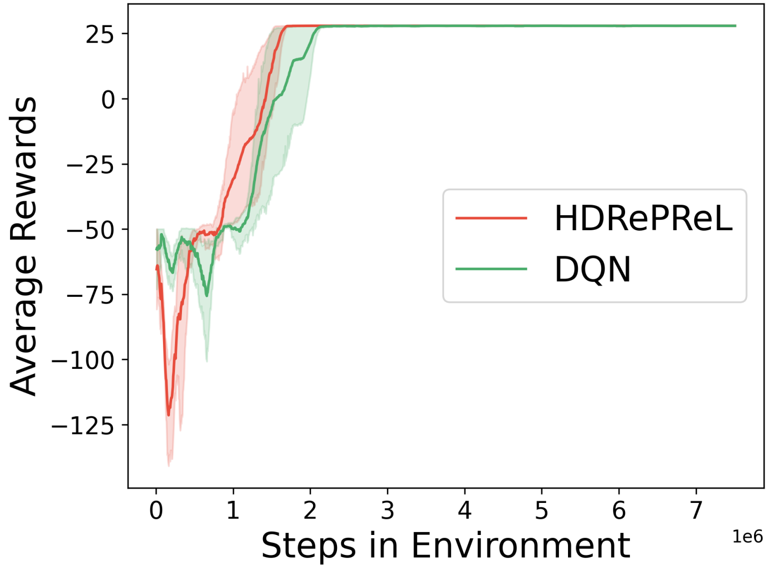


- Initialize buffers
- Get high level plan
- For each subgoal
  - *Loop till the subgoal is achieved or # steps exceeds*
    - Get the abstract state
    - Get the policy for that subgoal
    - Take a step and observe reward, next state
    - Add <S, A, R, S> to the buffer
- Update the subgoal policy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

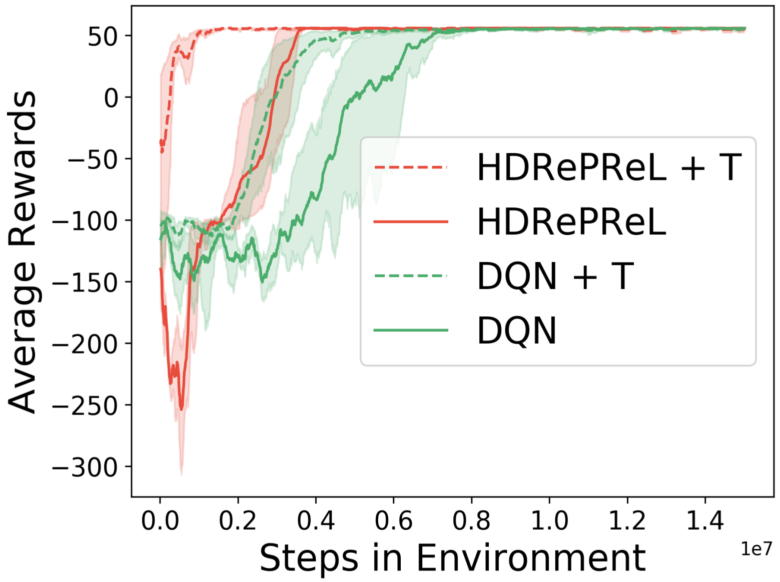
Experiments

- Sample efficiency
- Generalization across objects

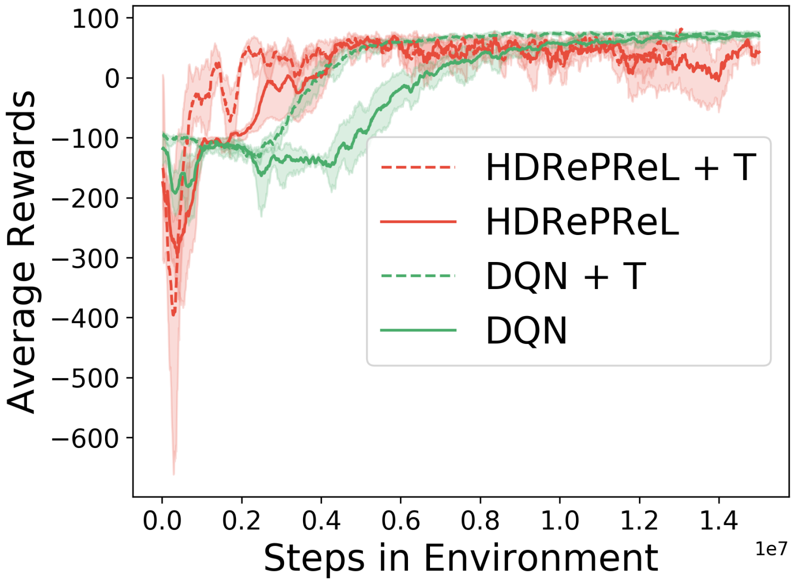




transport one passenger



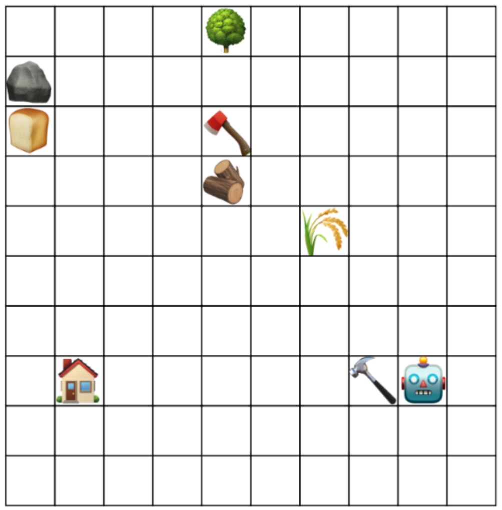
Transport two passengers

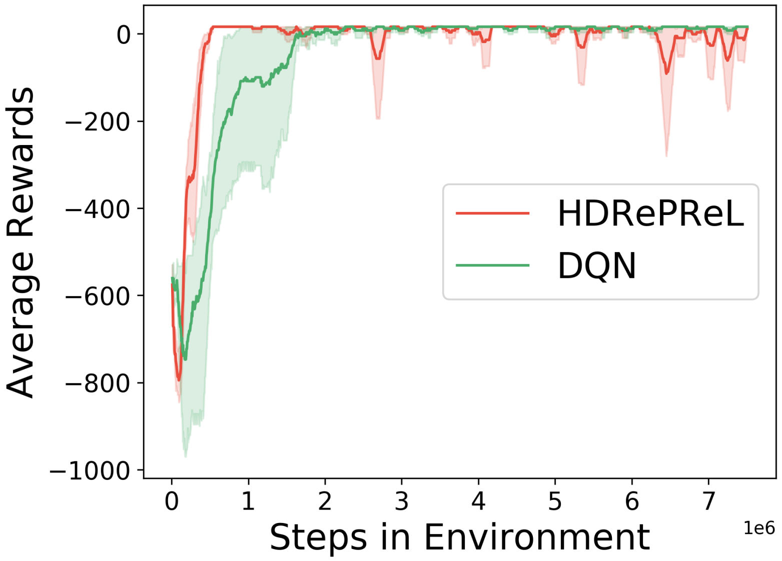


Transport three passengers

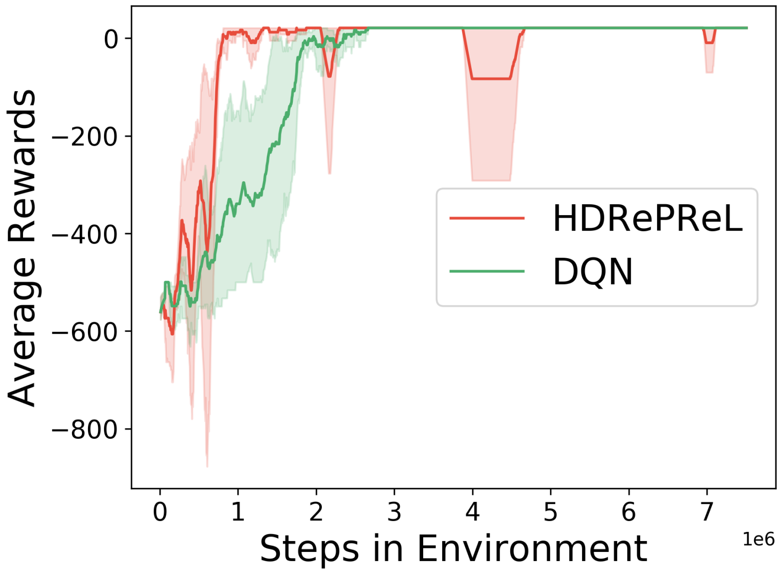
Experiments

- Sample efficiency
- Generalization across objects

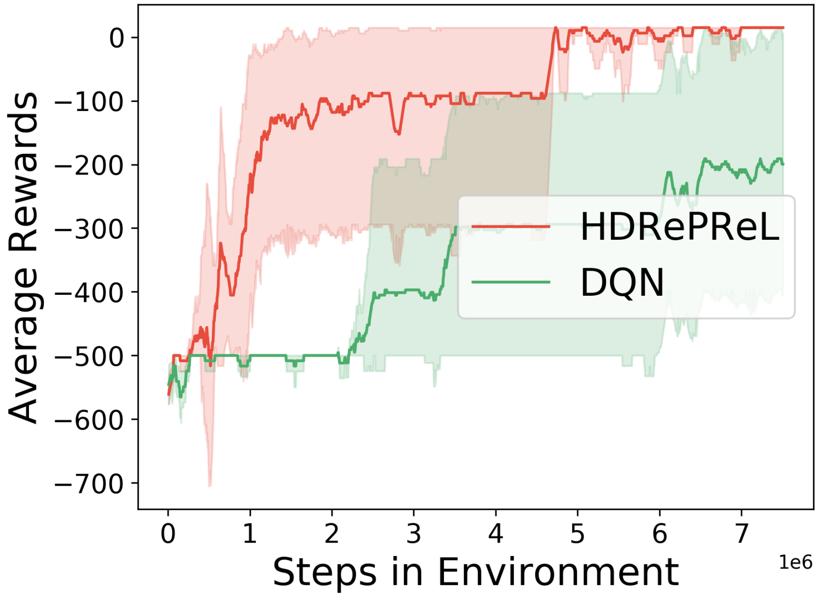




Make Bread



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

StARLinG LAB

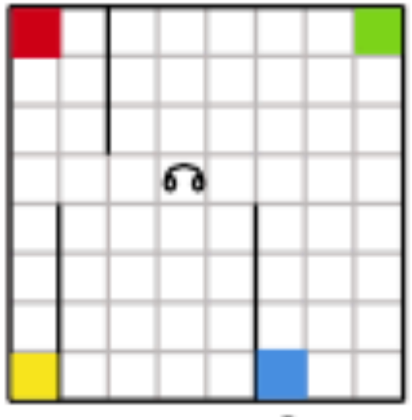
starling.utdallas.edu
harshakokel.com

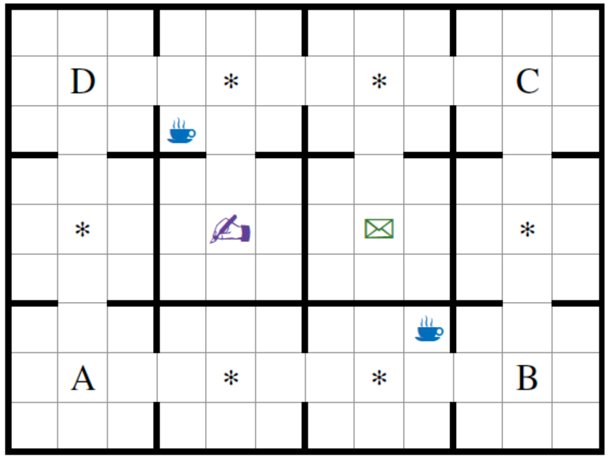
@starling_lab
@harsha_kokel

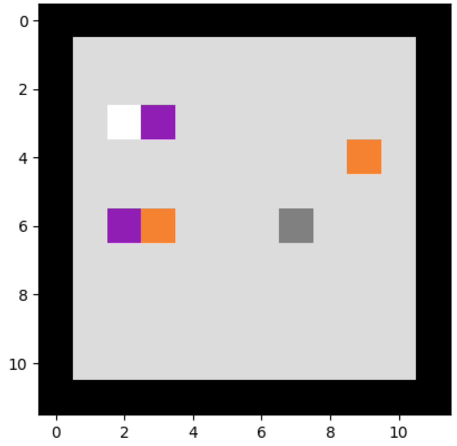
HARSHA STOPPED HERE

Experiments

- Domains
  - *Office World*
  - *Craft World*
  - *Relational Taxi*
  - *Relational Box World*
- Baselines
  - *HRL (options framework)*
  - *TRL (Taskable RL, Illanes et al. 2020)*

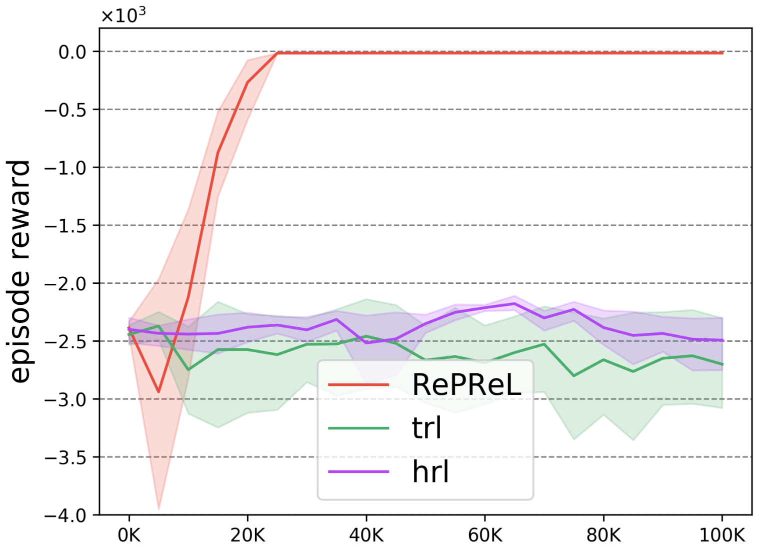




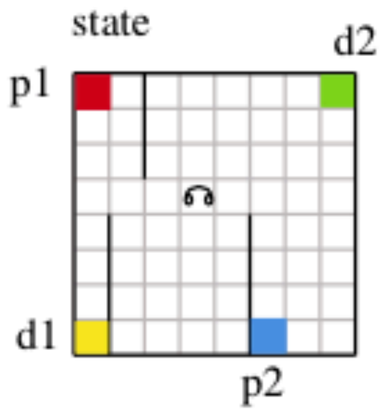


Experiments

- Sample efficiency
- Transfer across task
- Generalization across objects



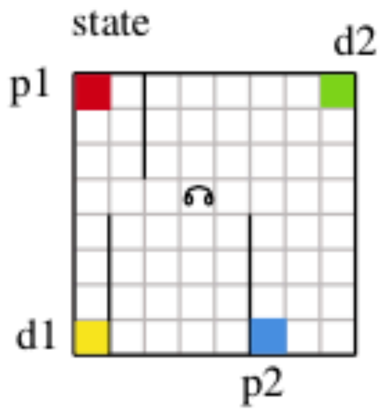
transport p1

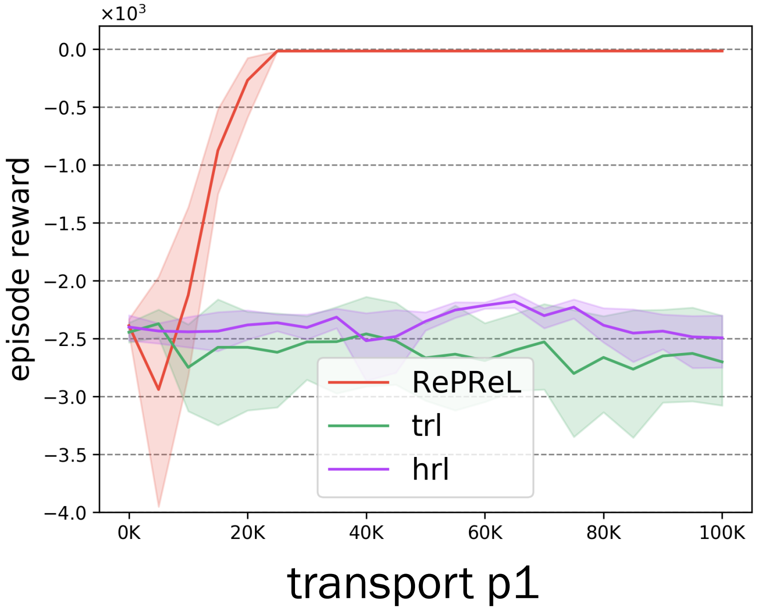


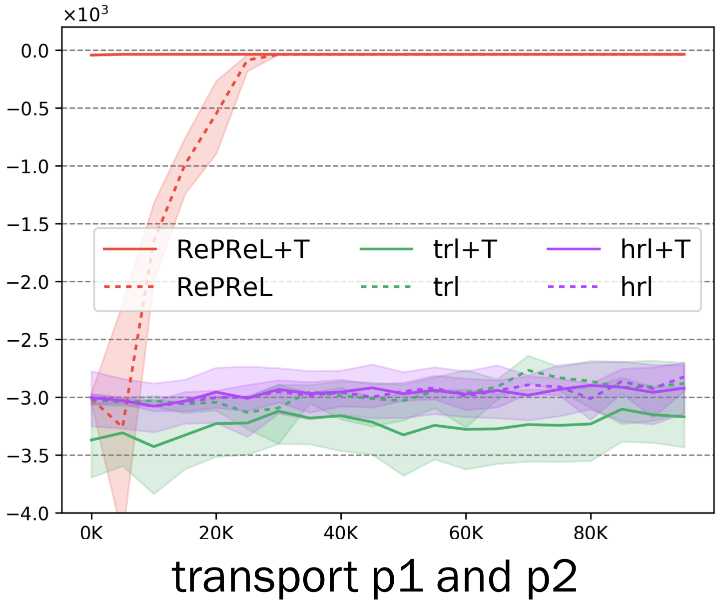
trl: Taskable RL (Illanes et al. ICAPS 2020)

Experiments

- Sample efficiency
- Transfer across task
- Generalization across objects

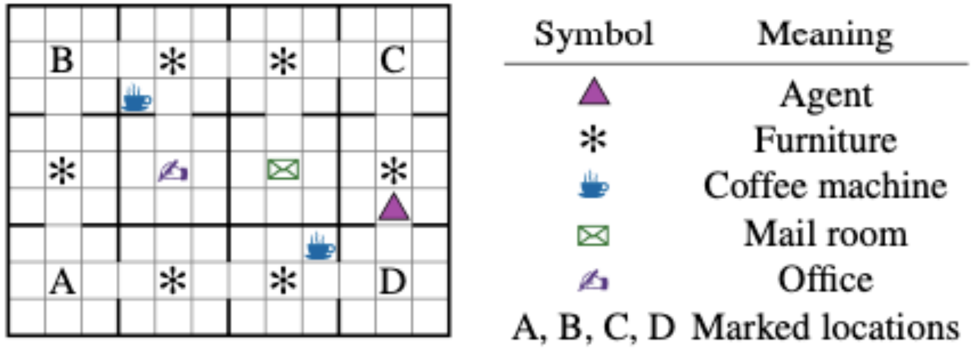


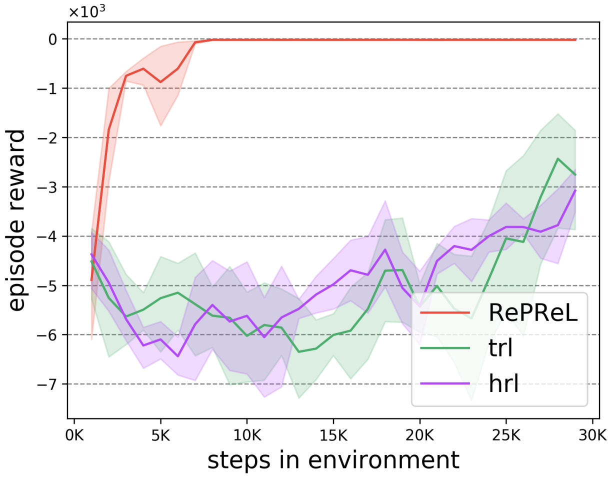




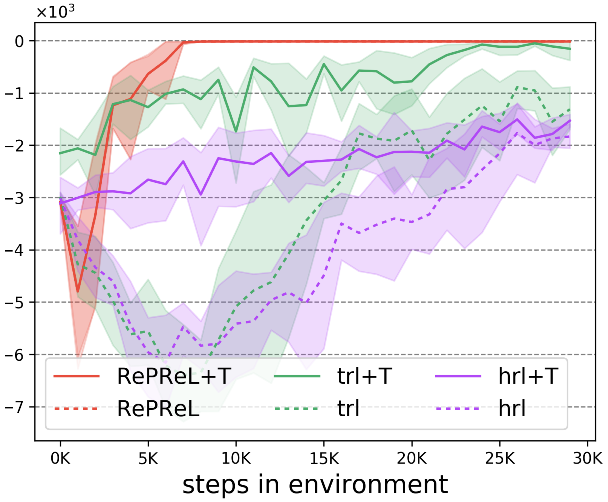
RePReL

Office World

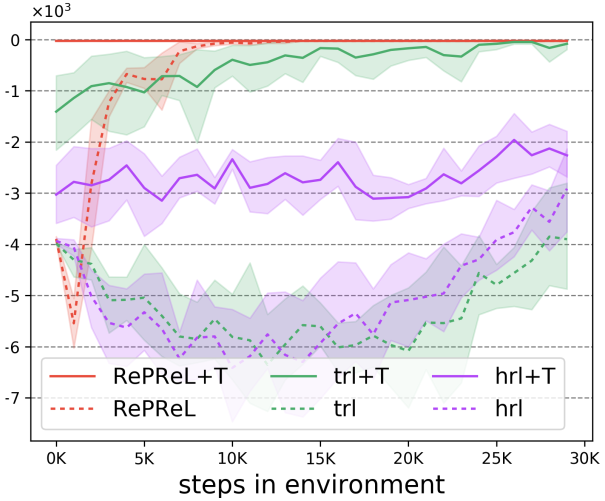




Deliver mail



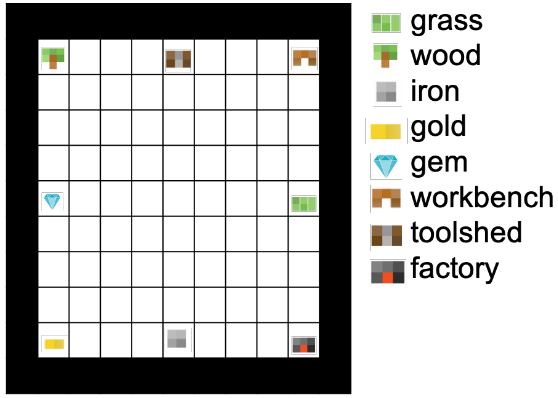
Deliver coffee

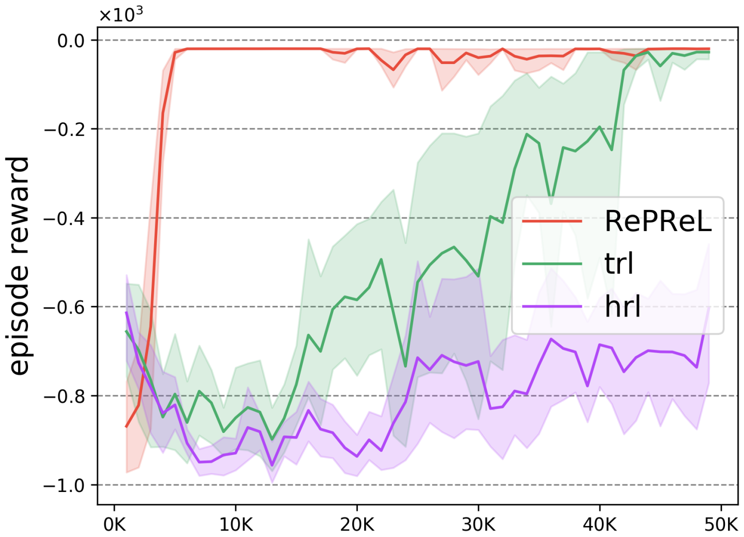


Deliver mail and coffee

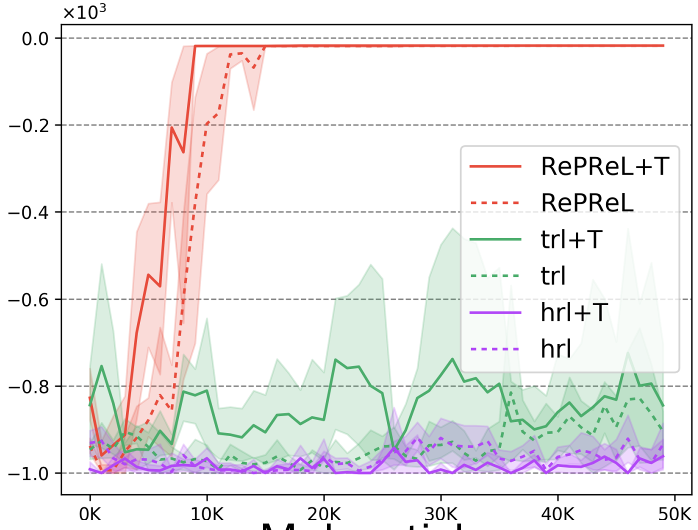
RePReL

CRAFT WORLD





Get wood and iron



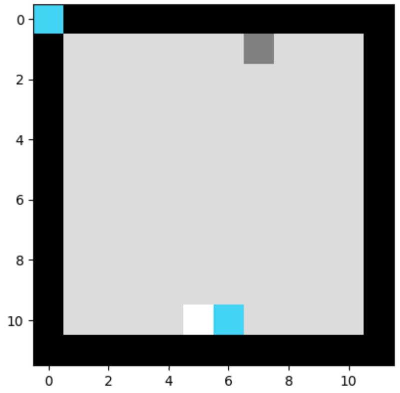
Make stick

(at workbench, needs wood)

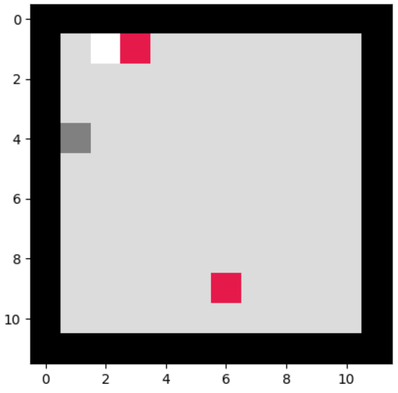
RePReL

Relational Box World

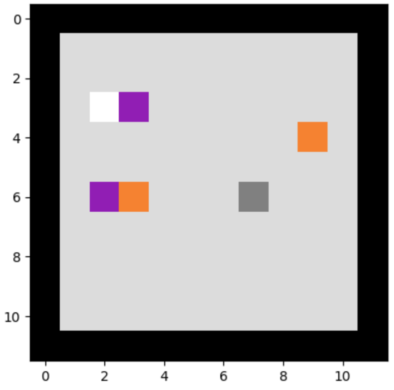
Generalization across objects



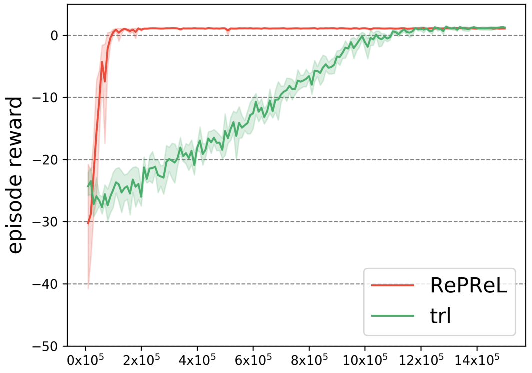
Open lock

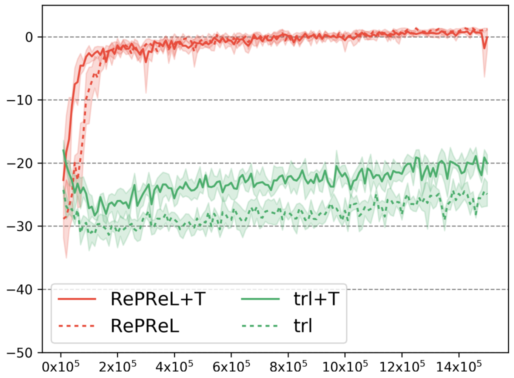


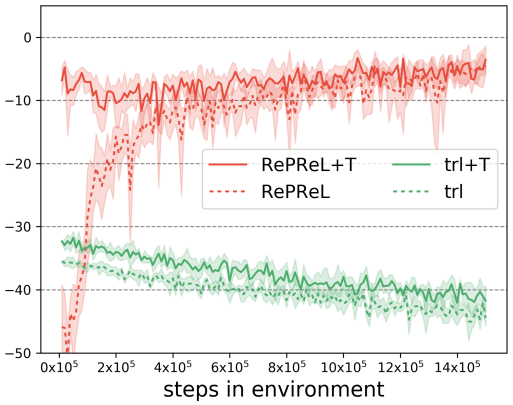
Collect key and open lock



Collect key and open 2 locks







For human-level general intelligence, the ability to detect compositional structure in the domain and form task-specific abstractions are necessary.

Other relevant work

- Learning the high-level planner [*Ludovico et al IJCLR 2021*]
- Modify the plan based on RL agents capability [Lyu et al AAAI 2019]
- Automating task termination condition [Lee et al 2021]
- Learning task-specific state representation [Abdulhai et al. 2021]
- Learning to plan and act simultaneously [Patra et al AI 2021]
- Improving Robot Navigation [Wöhlke et al. ICRA 2021]
- Extending the RePReL framework to Deep RL setting (under prep)