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

—

Harsha Kokel

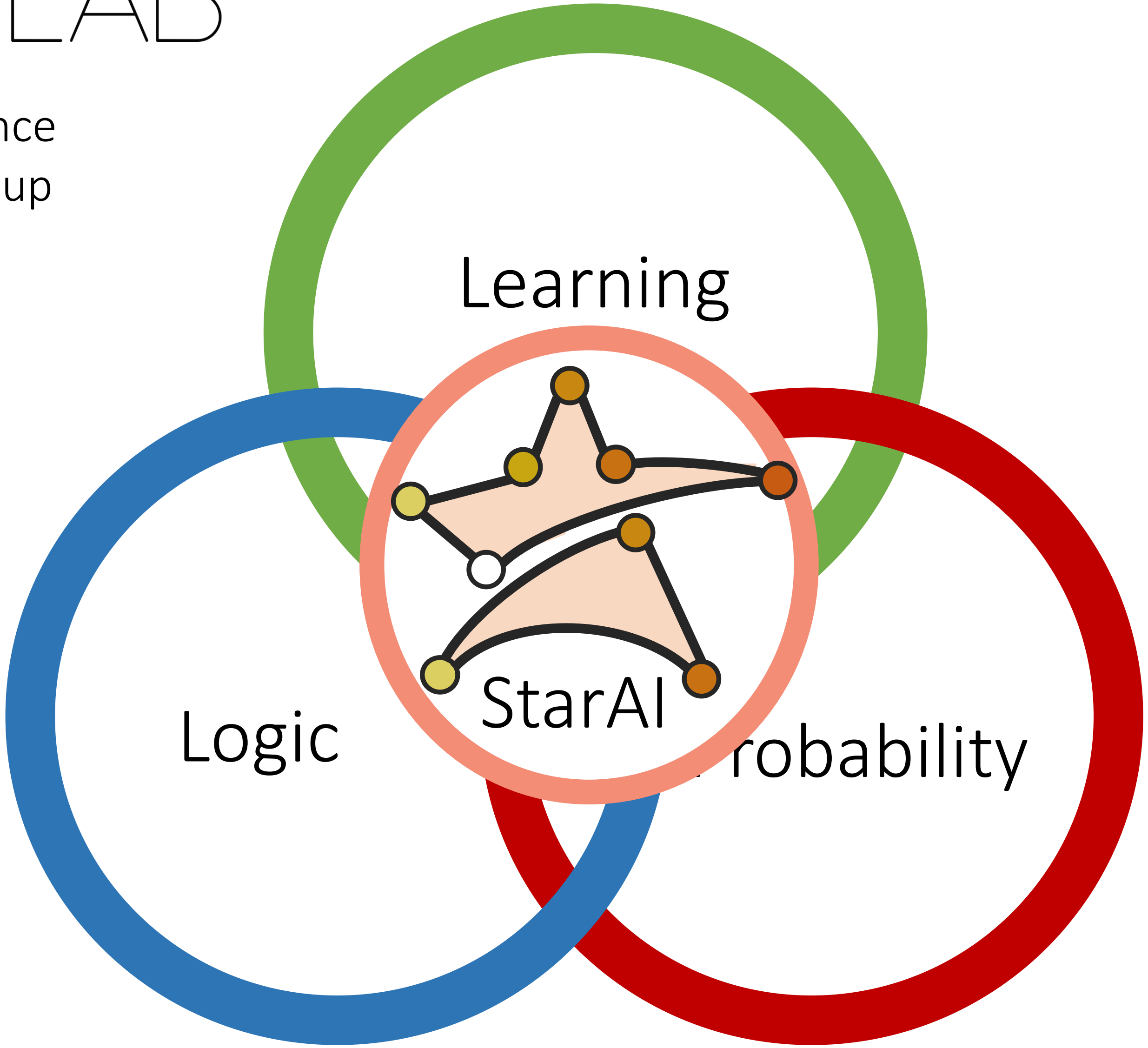
Dynamic probabilistic logic models for effective task-specific abstractions in RL

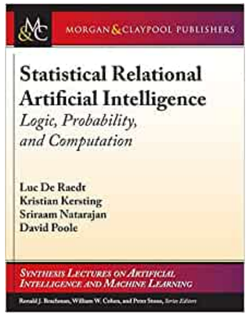
—

Harsha Kokel
Nikhilesh Prabhakar
Sriraam Natarajan
Arjun Manoharan
Balaraman Ravindran
Erik Blasch
Prasad Tadepalli

With support from DARPA, AFOSR, ARO, NSF & USDA-NIFA, and RBCDSAI

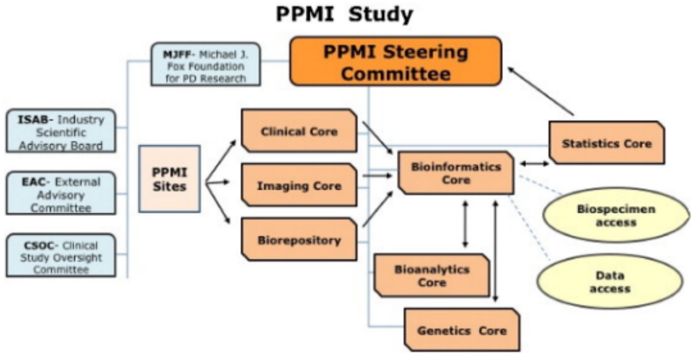
Statistical Artificial Intelligence
and Relational Learning Group



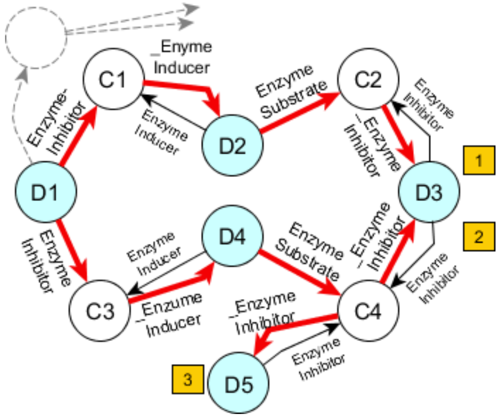


Raedt et al. 2016; Raedt et al. 2020

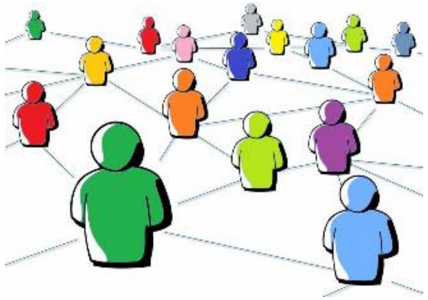
StARLinG LAB



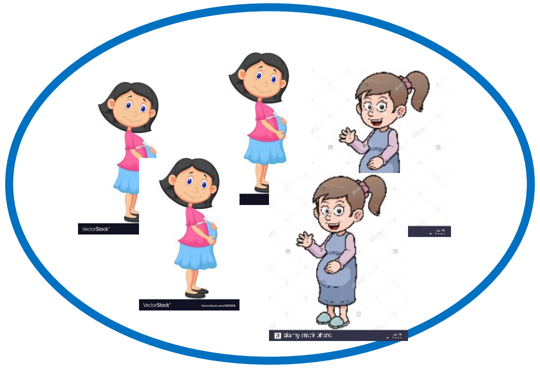
Parkinson's disease prediction



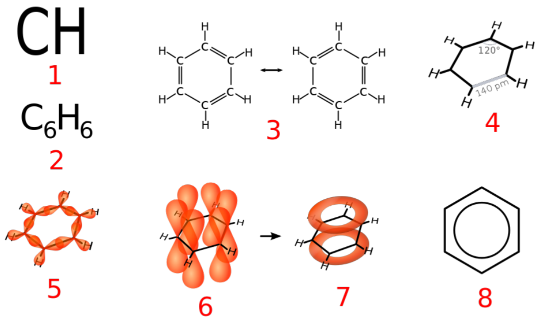
Drug-Drug Interactions



Social Networks



Cohort of Pregnant Women (nuMoM2b)

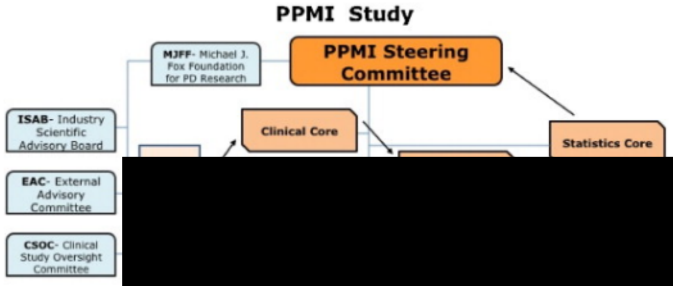


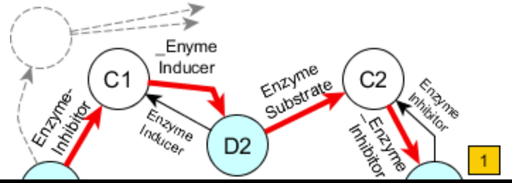
Chemical Entities of Biological Interest (ChEBI)

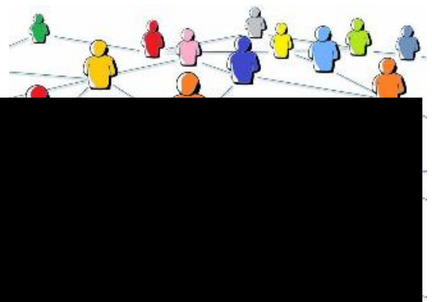
Collaborative Problem Solving

Dhami et al. 2017; Dhami et al. 2022; Karanam et al. 2022; Mathur et al. 2023; Dhami et al. 2017; Das et al. 2022; Ramanan et al. 2021; Kaur et al. 2020; Kokel et al. 2021

StARLinG LAB

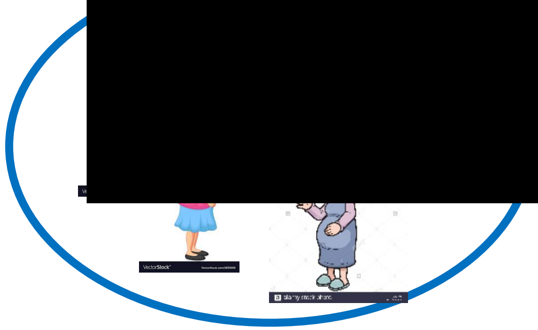




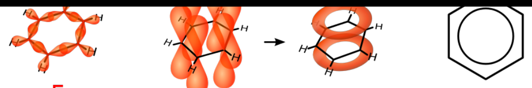


Park

Relational and Human-Allied



Cohort of Pregnant Women (nuMoM2b)

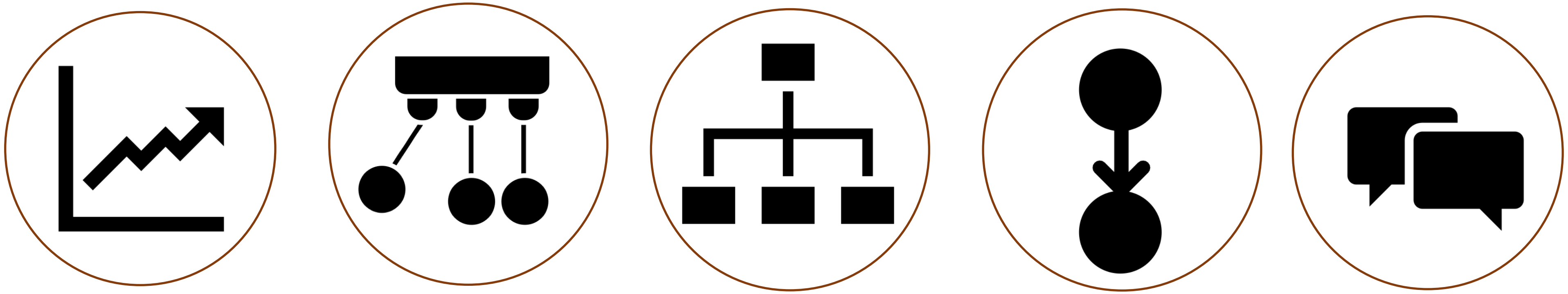


Chemical Entities of Biological Interest (ChEBI)

Collaborative Problem Solving

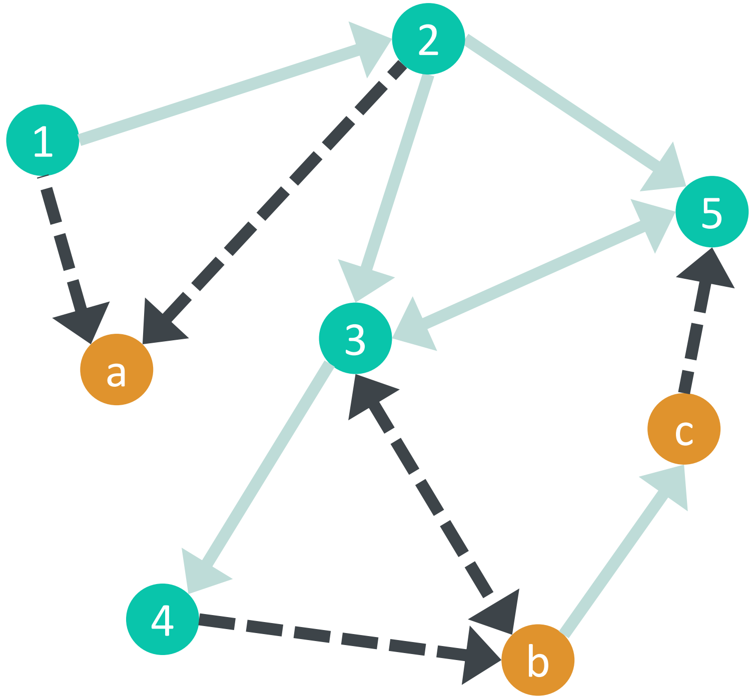
Dhami et al. 2017; Dhami et al. 2022; Karanam et al. 2022; Mathur et al. 2023; Dhami et al. 2017; Das et al. 2022; Ramanan et al. 2021; Kaur et al. 2020; Kokel et al. 2021

How to facilitate generalizable, effective and efficient learning with human guidance?

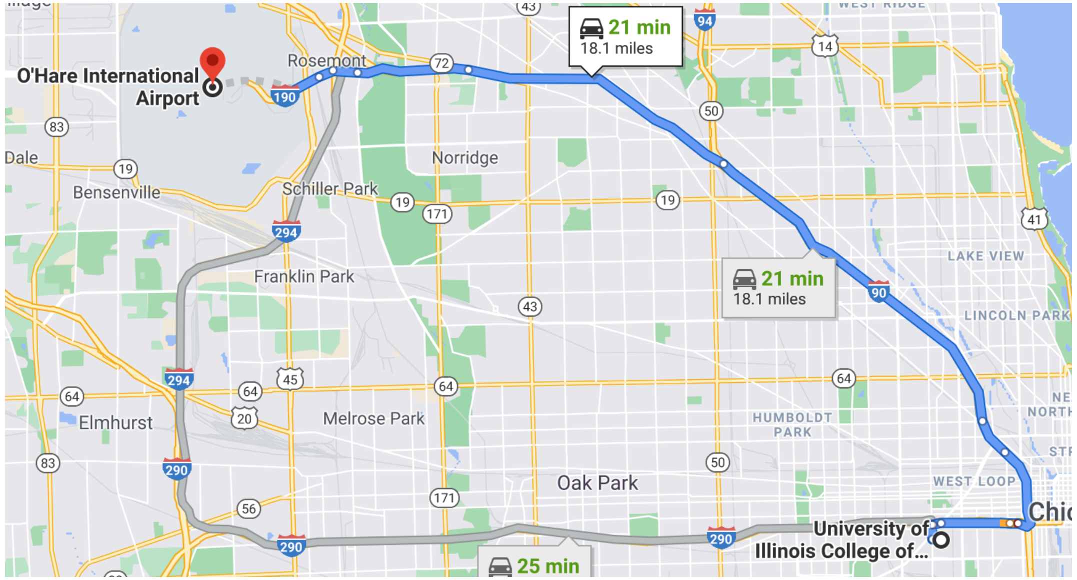


Relational domains

Non-IID domains with varying # objects and heterogeneous relations.



Abstract Representations



Planning

Execution

Konidaris, G., 2019; Li et al 2006

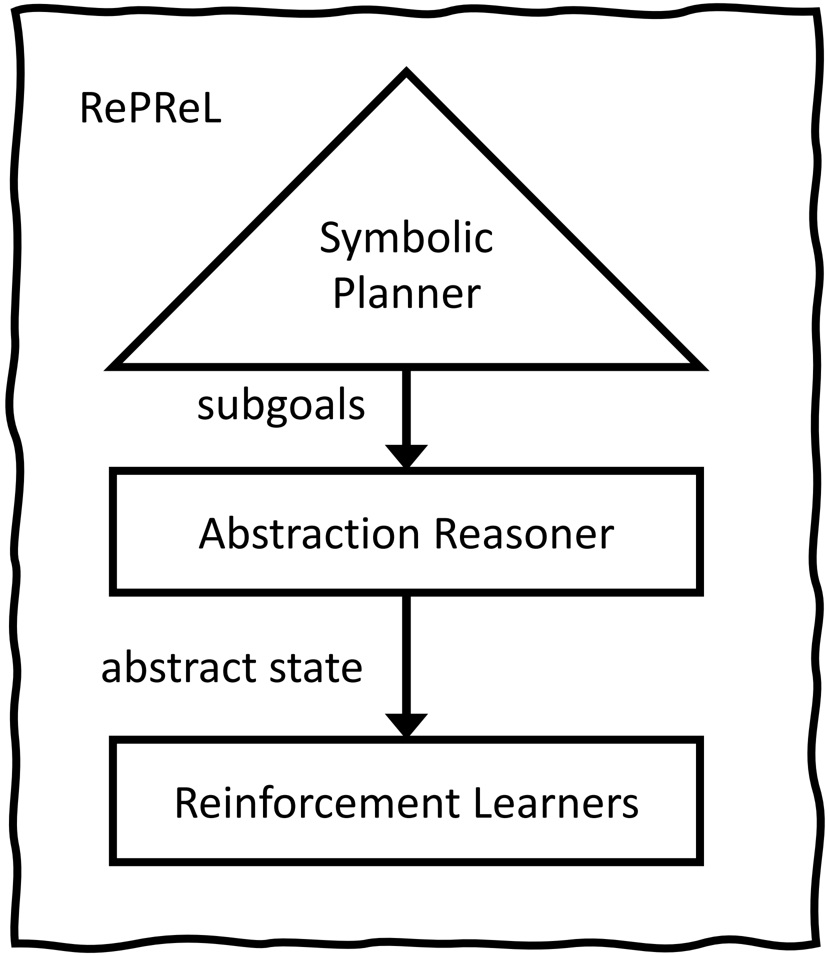
Given: Relational sequential decision-making domain

To do: Learn an efficient agent that

- is compositional
- can handle varying # of objects
- can generalize to different tasks
- can support task-specific representations
- can handle multi-modal data

RePReL

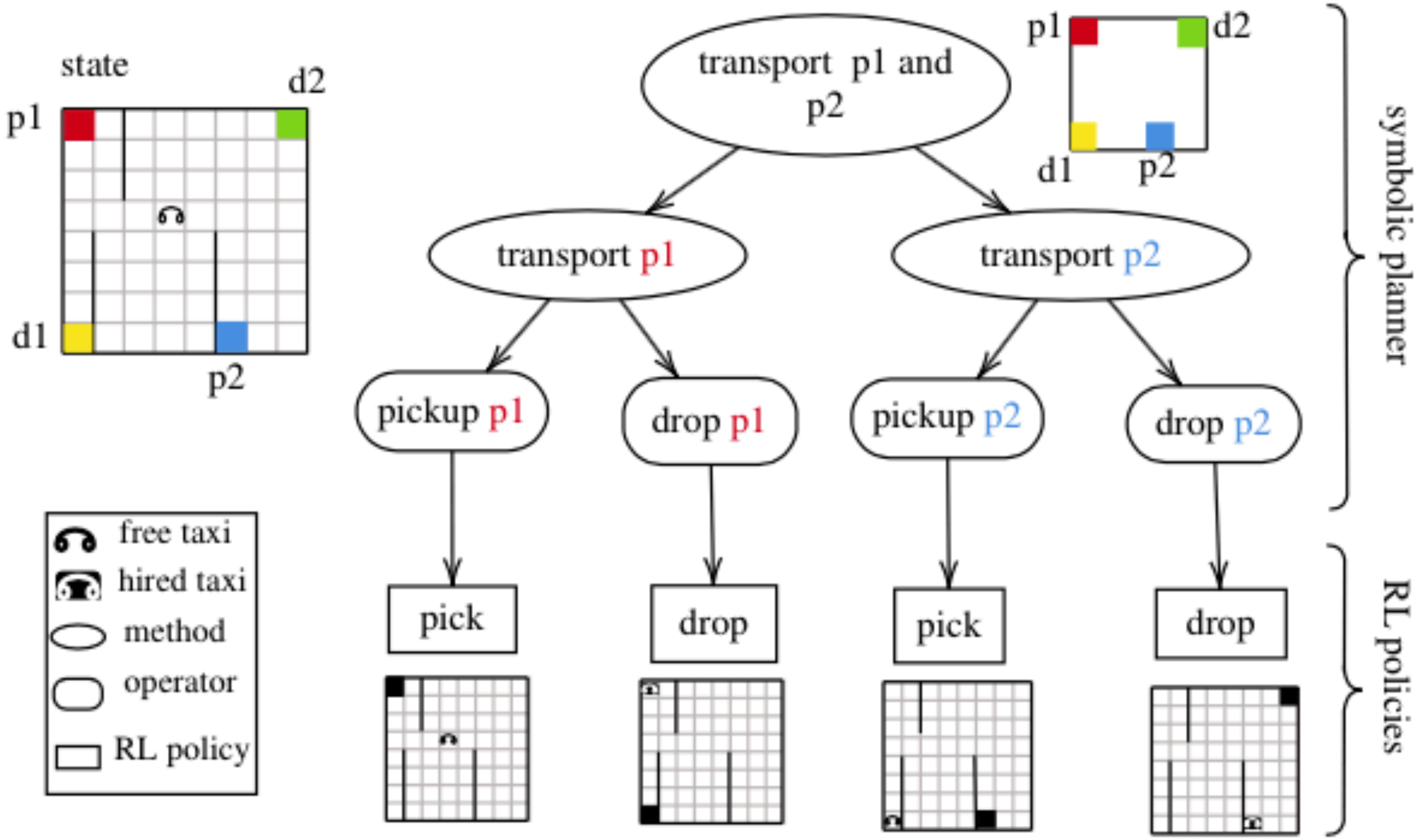
- Plan the sequence of high level subgoals and learn to execute each subgoal at lower level



Grounds and Kudenko 2008; Yang et al. 2018; Jiang et al. 2019; Eppe et al. 2019; Illanes et al. 2020; Lee et al. 2020; Mitchener et al. 2022; Lyu et al. 2019; Goel et al. 2022; Planning and RL workshop

RePReL

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



Dietterich 1999

RePReL

Definition 3. *The subgoal RMDP M_o for each operator o is defined by the tuple $\langle S, A, P_o, R_o, \gamma \rangle$ consisting of states S, actions A, transition function P_o, reward function R_o, and discount factor γ. State and Actions remain same as the original RMDP. The reward function R_o and transition probability distribution function P_o are defined as follows:*

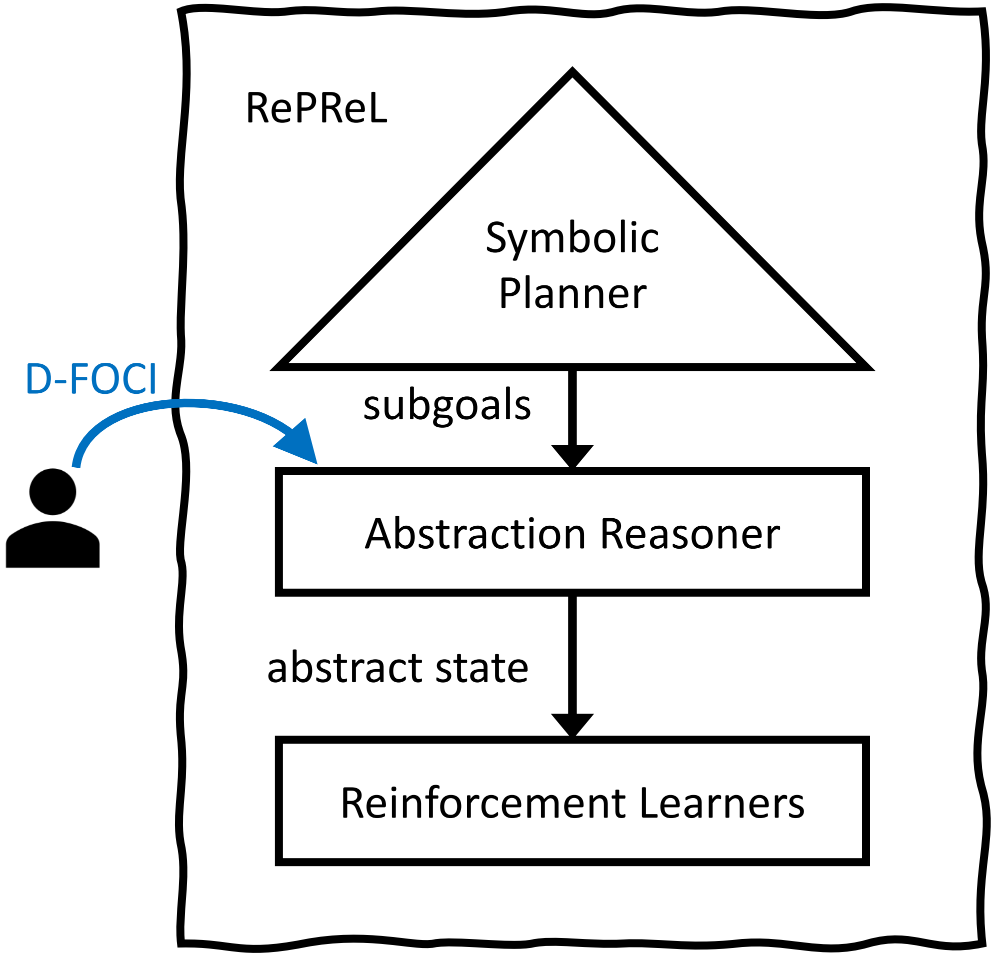
$$R_o(s,a,s') = \begin{cases} t_R + R(s,a,s') & \textit{if } s' \in \beta(o) \textit{ and } s \notin \beta(o) \\ 0 & \textit{if } s' \in \beta(o) \textit{ and } s \in \beta(o) \\ R(s,a,s') & \textit{otherwise} \end{cases}$$

$$P_o(s,a,s') = \begin{cases} 0 & \textit{if } s \in \beta(o) \textit{ and } s' \notin \beta(o) \\ 1 & \textit{if } s \in \beta(o) \textit{ and } s' \in \beta(o) \\ P(s,a,s') & \textit{otherwise} \end{cases}$$

with $R(s,a,s')$ indicating the reward function from the original GRMDP definition. t_R is a fixed terminal reward.

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 (D-FOCI) statements to obtain task-specific abstract representations



Kokel et al ICAPS 2021

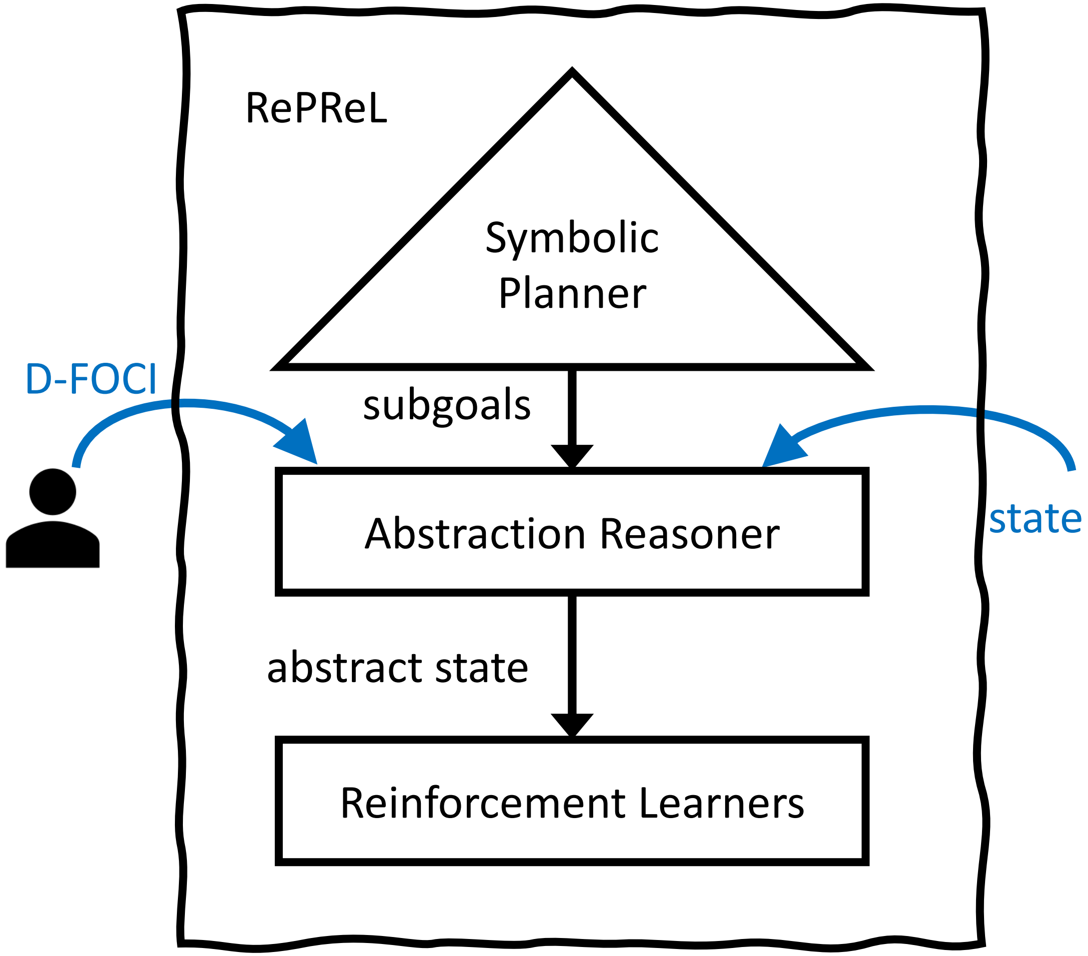
D-FOCI

First Order Conditional Influence (FOCI) statements

$$\text{if } \langle condition \rangle$$
$$\text{then } \langle influent \rangle \text{ QINF } \langle resultant \rangle$$

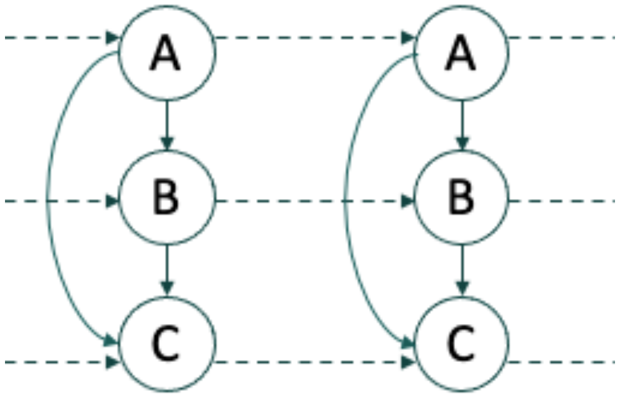
Dynamic FOCI statements

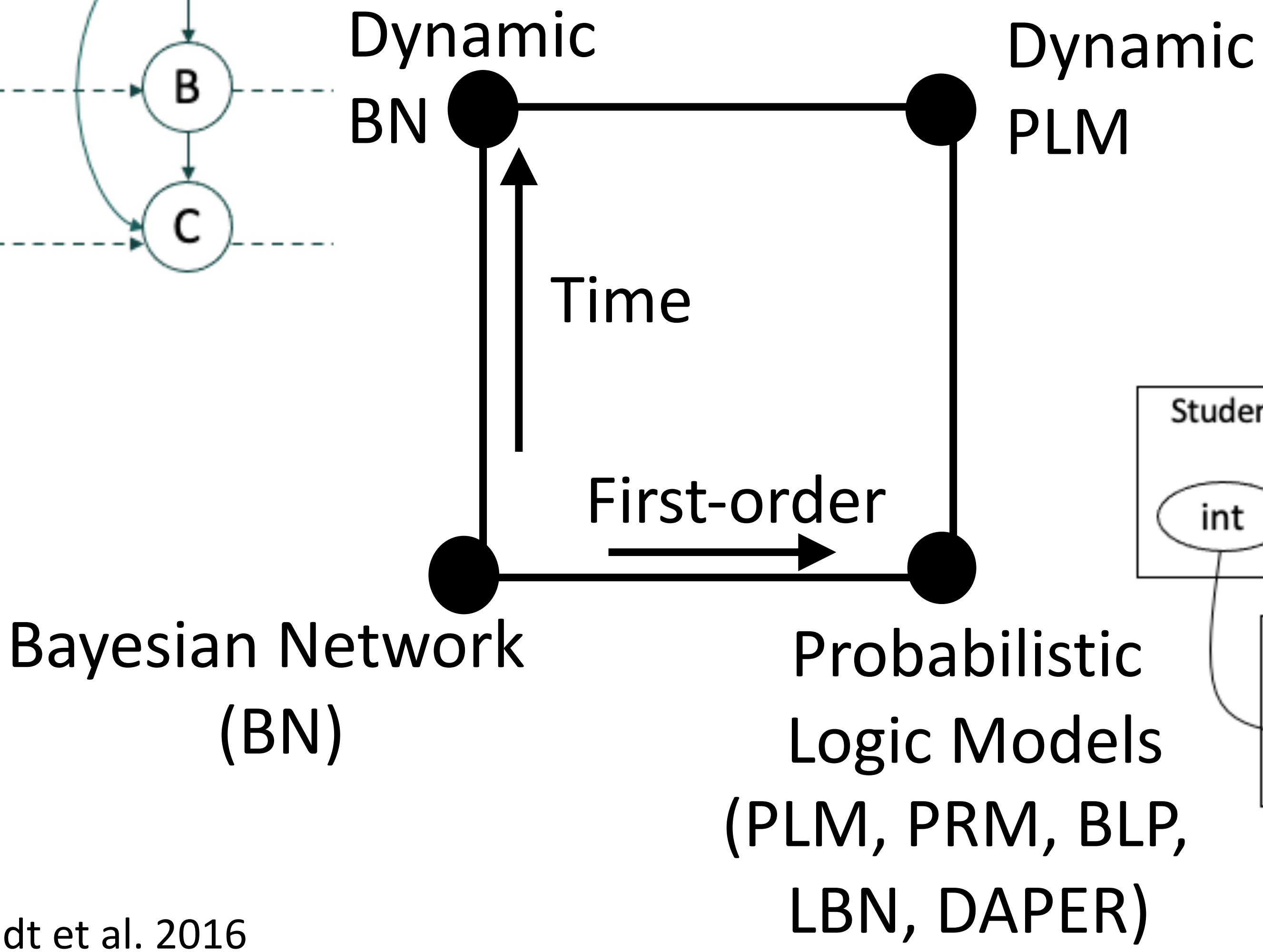
$$[\text{subgoal:}] \;\; \langle influent \rangle \xrightarrow{[+1]} \langle resultant \rangle$$

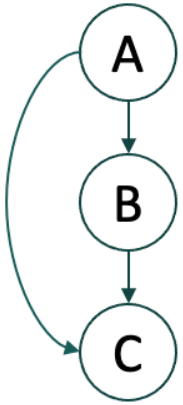


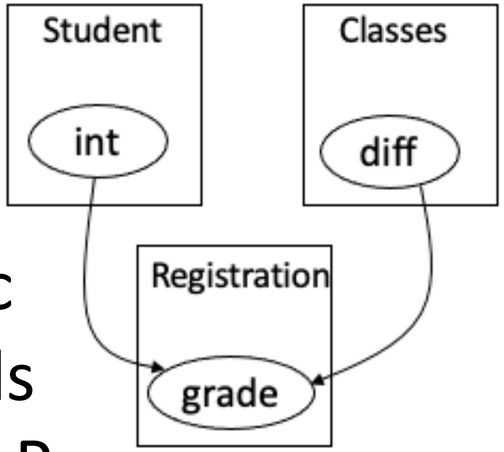
Natarajan, Tadepalli, Dietterich, and Fern 2008

D-FOCI as Dynamic PLMs









Koller and Friedman 2009;
Getoor and tasker 2007; Raedt et al. 2016

D-FOCI example

$$\{\texttt{action}, \texttt{taxi\_at}(X)\} \xrightarrow{+1} \texttt{taxi\_at}(X) \quad (3a)$$

$$\texttt{pick}(P) : \{\texttt{action}, \texttt{taxi\_at}(X), \texttt{at}(P, Y), \texttt{in\_taxi}(P)\} \xrightarrow{+1} \texttt{in\_taxi}(P) \quad (3b)$$

$$\texttt{pick}(P) : \{\texttt{in\_taxi}(P)\} \longrightarrow \mathit{Reward} \quad (3c)$$

$$\texttt{drop}(P) : \{\texttt{at\_dest}(P)\} \longrightarrow \mathit{Reward} \quad (3d)$$

$$\texttt{drop}(P) : \{\texttt{at}(P, X), \texttt{dest}(P, D), \texttt{at\_dest}(P)\} \longrightarrow \texttt{at\_dest}(P) \quad (3e)$$

$$\texttt{drop}(P) : \{\texttt{action}, \texttt{taxi\_at}(X), \texttt{at}(P, Y), \texttt{in\_taxi}(P)\} \xrightarrow{+1} \texttt{at}(P, K) \quad (3f)$$

Abstraction

Given:

a. D-FOCI statements from Equation 3

b. state $s = \{$ at(p1,r), taxi_at(l3), dest(p1,d1), ¬at_dest(p1) ¬in_taxi(p1), at(p2,b), ¬at_dest(p2), ¬in_taxi(p2)$\}$

c. grounded option$o\theta$: pick(P) $\{P/\text{p1}\}$

Output: A set of relevant state literals: $\hat{s}$

Depth 1 unrolling:

1. Find a substitution that grounds relevant D-FOCI statements that have reward on RHS

   pick(p1): in_taxi(p1) $\longrightarrow$ *Reward*

   $\theta = \{\text{P/p1}\}$

2. Collect LHS in relevant literals set $\hat{s}$

   $\hat{s} \leftarrow \{$in_taxi(p1)$\}$

Depth 2 unrolling:

1. Find a substitution that grounds relevant D-FOCI statements that have a relevant literal on RHS

   pick(P): $\{$ action, taxi_at(l3), at(p1, r), in_taxi(p1) $\}$ $\longrightarrow$ in_taxi(p1)

   $\theta = \{\text{P/p1, X/l3, Y/r}\}$

2. Collect LHS in set $\hat{s}$

   $\hat{s} \leftarrow \{$in_taxi(p1), action, taxi_at(l3), at(p1, r)$\}$

Depth 3 unrolling:

1. Ground applicable D-FOCI statements that have a relevant literal ($\hat{s}$) on RHS

   $\{$action, taxi_at(l3) $\}$ $\xrightarrow{+1}$ taxi_at(l3)

   pick(p1): $\{$ action, taxi_at(l3), at(p1, r), in_taxi(p1) $\}$ $\longrightarrow$ in_taxi(p1)

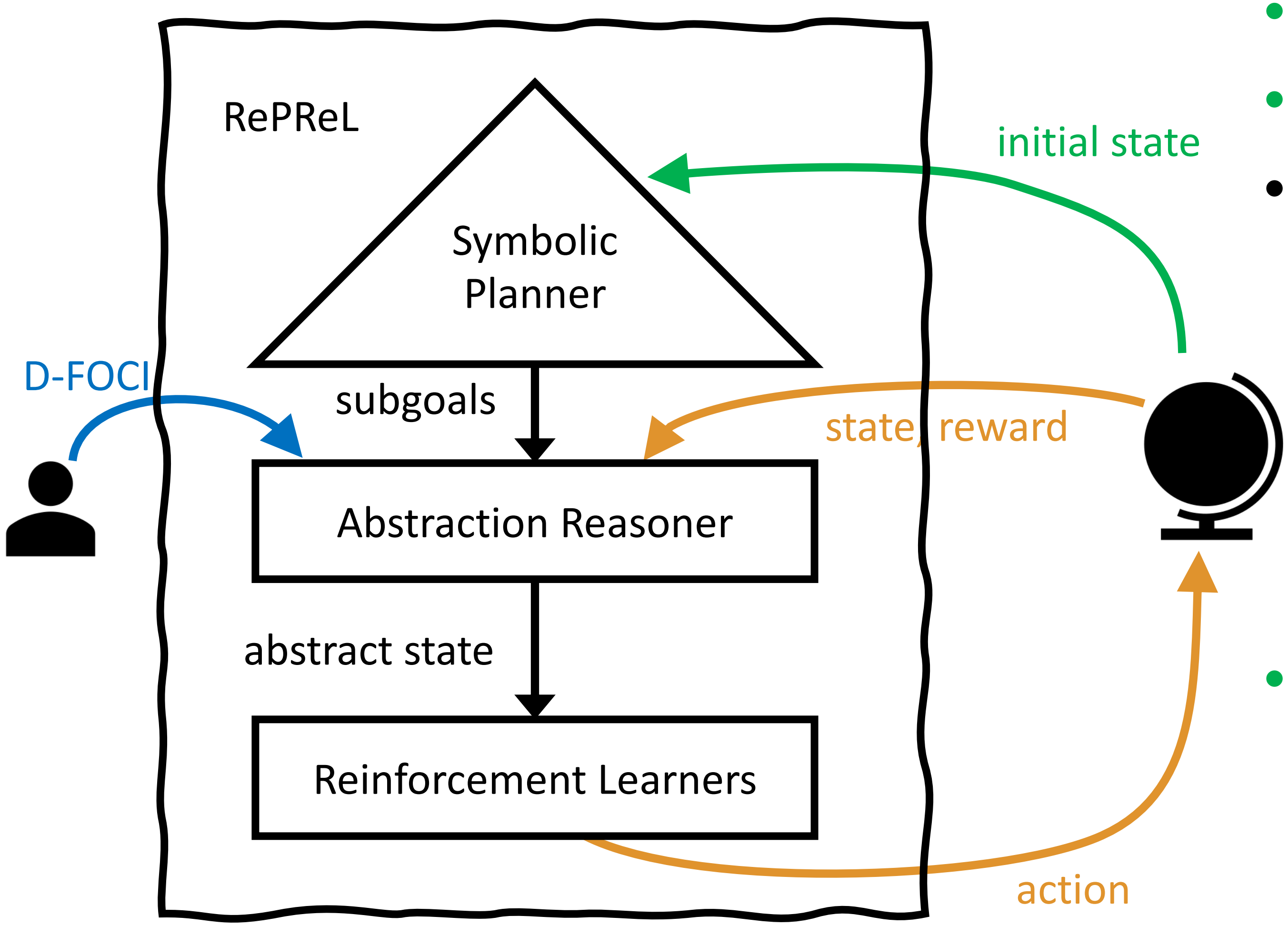
   $\theta = \{\text{P/p1, X/l3, Y/r}\}$

2. Collect LHS in set $\hat{s}$

   $\hat{s} \leftarrow \{$in_taxi(p1), action, taxi_at(l3), at(p1, r)$\}$

recursive grounding and unrolling process

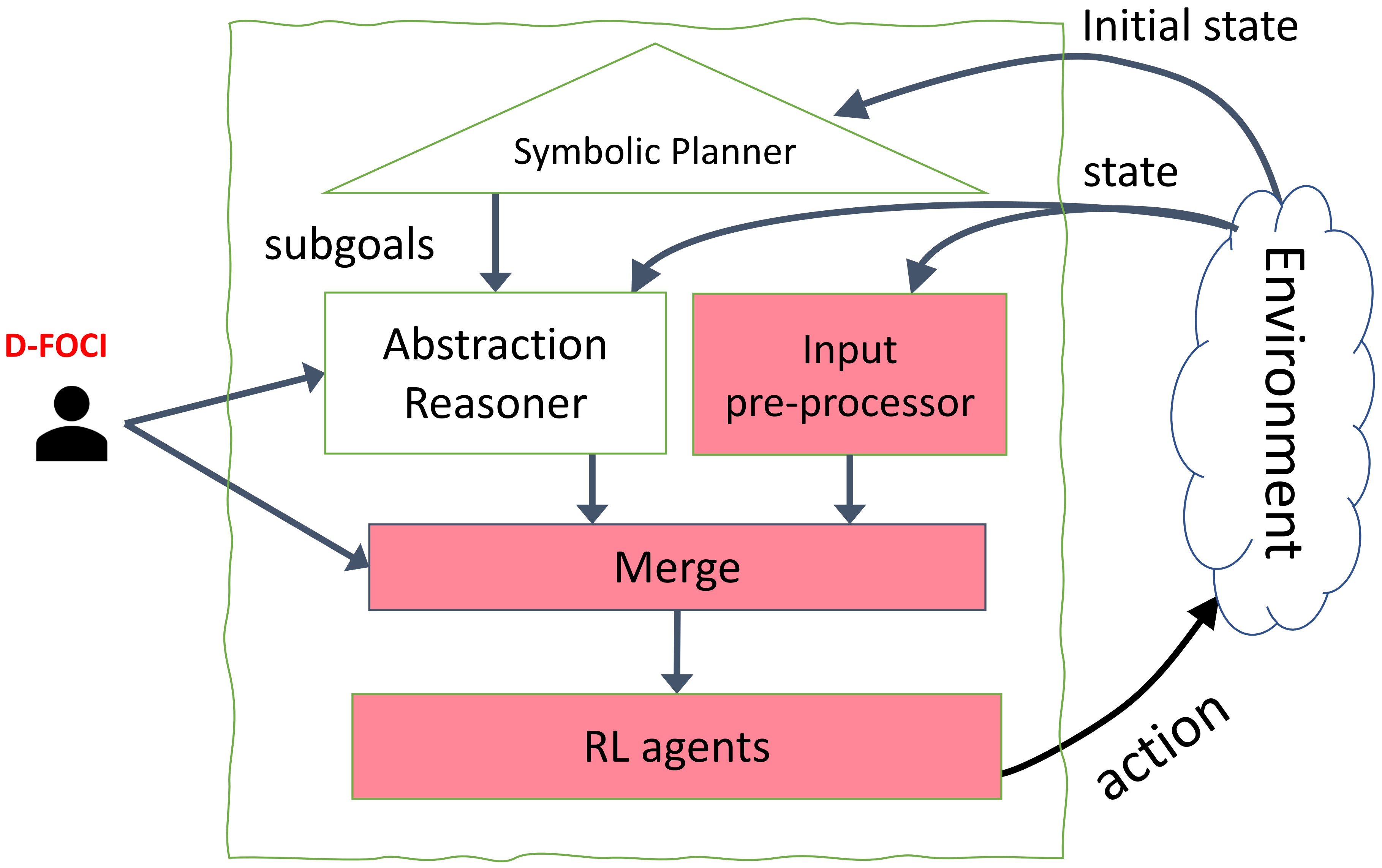
RePReL Learning



- 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

Kokel et al. 2021a; Kokel et al. 2021b

Hybrid Deep RePReL



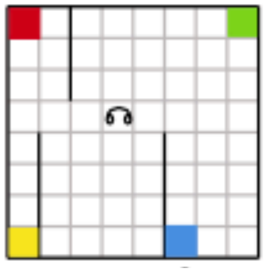
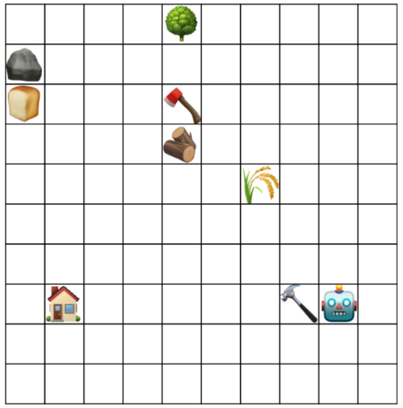
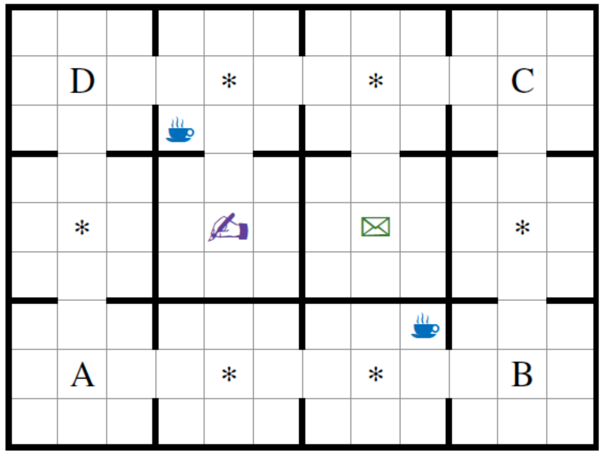
Manhaeve et al. 2018; Kokel et al. 2022

Given: Relational sequential decision-making domain

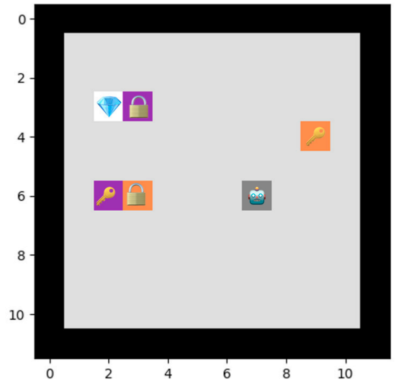
To do: Learn an efficient agent that

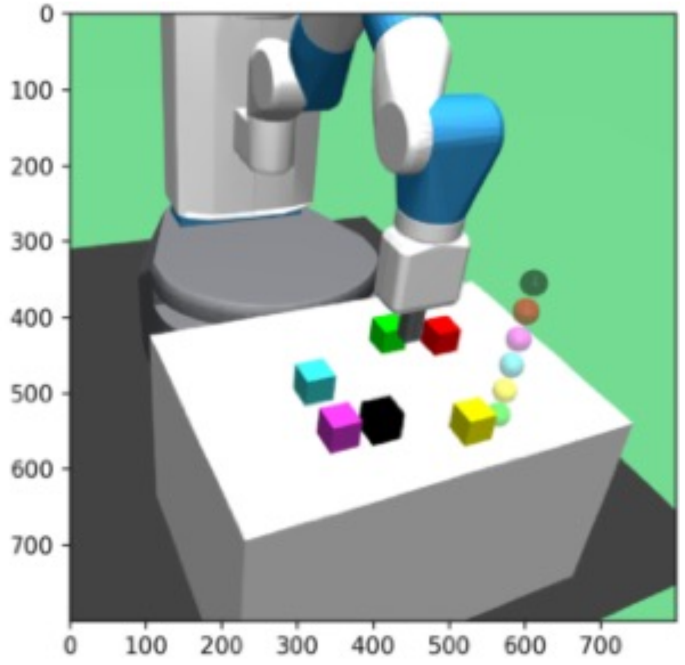
- is compositional
- can handle varying # of objects
- can generalize to different tasks
- can support task-specific representations
- can handle multi-modal data

Experiments

- Domains
  - Office World
  - Craft World
  - Relational Taxi
  - Relational Box World
  - Fetch Pick and Place

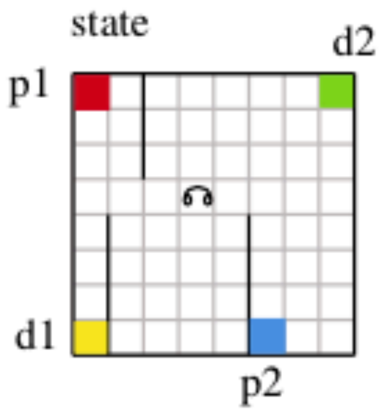


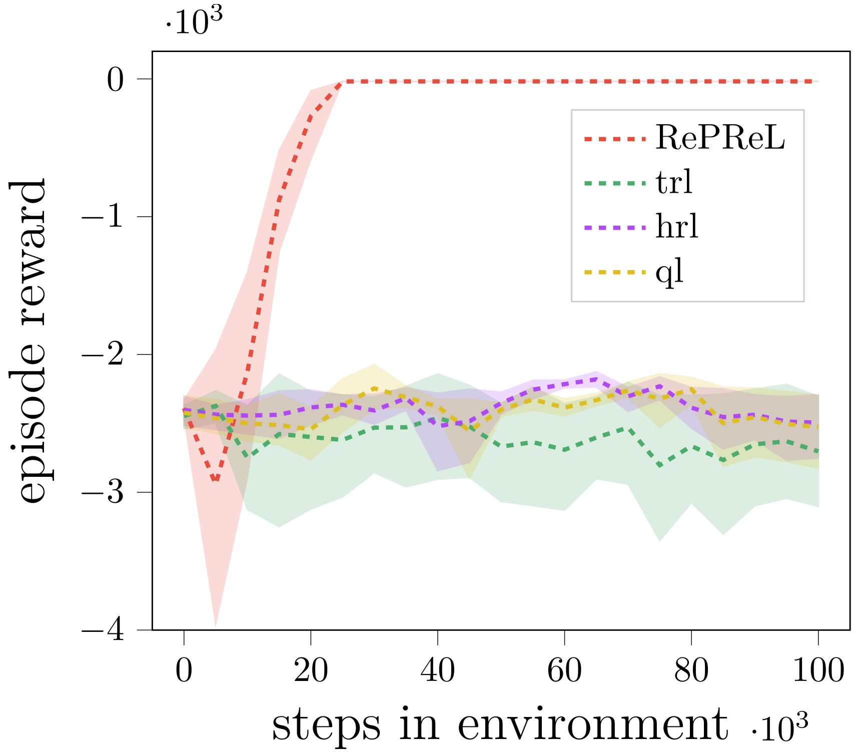


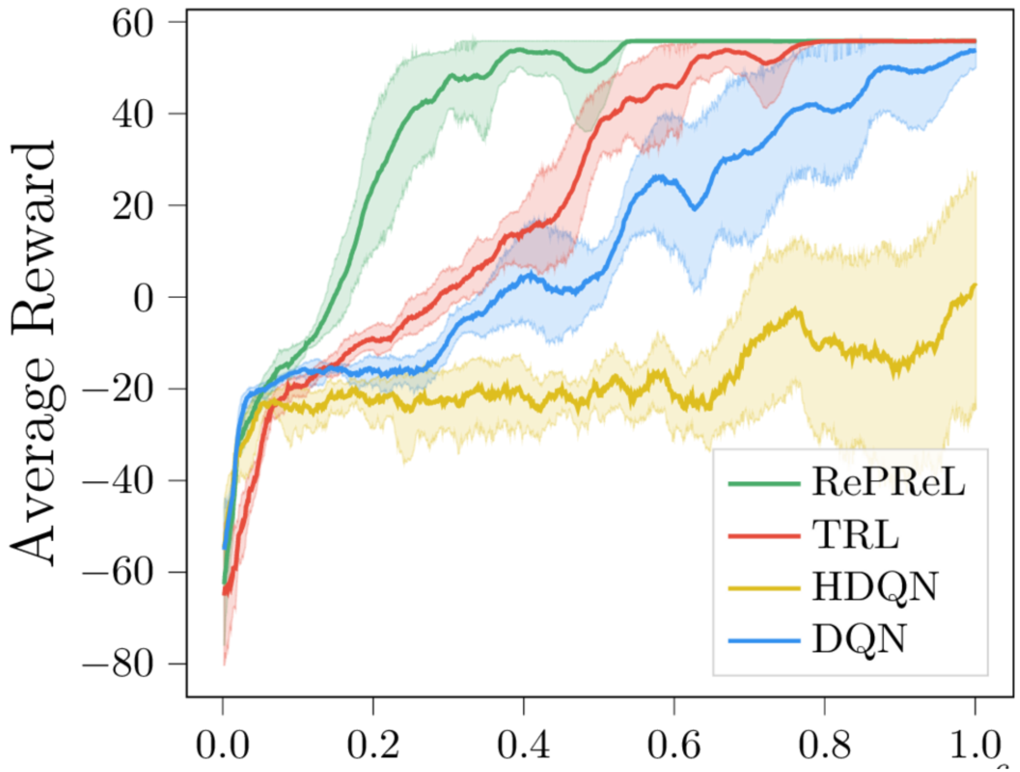
- Baselines
  - Tabular RL
  - Deep RL (DDQN, PPO, SAC)
  - Hierarchical RL (options framework)
  - Planner + RL (Taskable RL)
  - Deep Relational RL (ReNN)

Illanes et al. 2020; Li et al 2020;

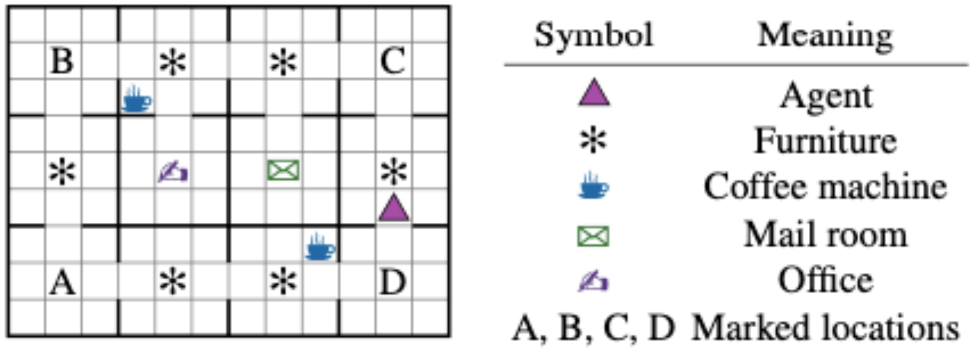
Sample Efficiency

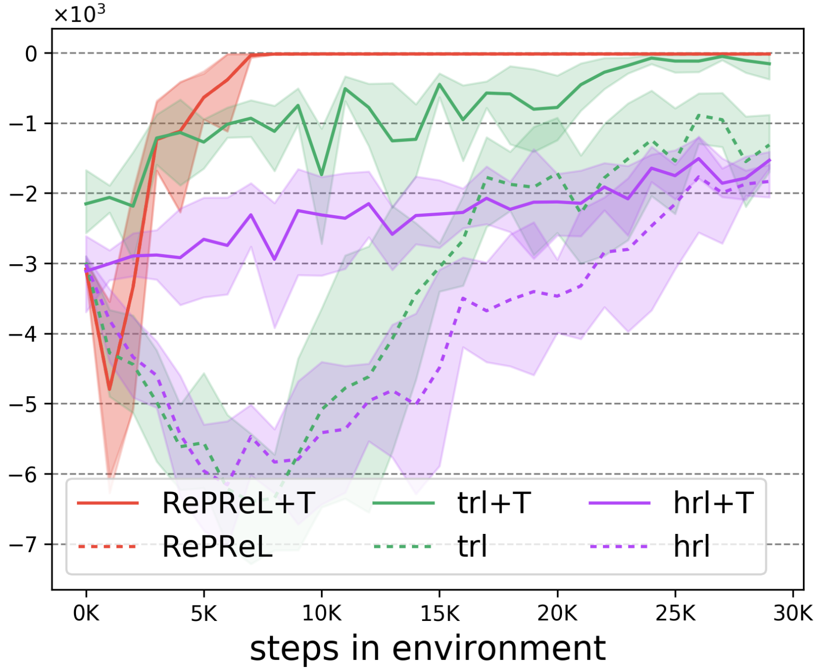


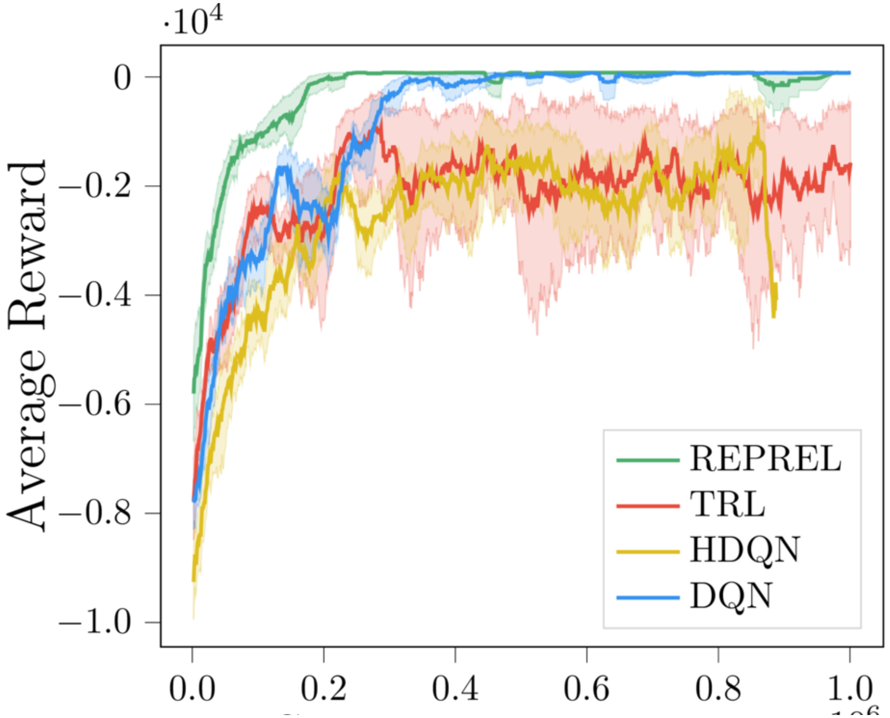




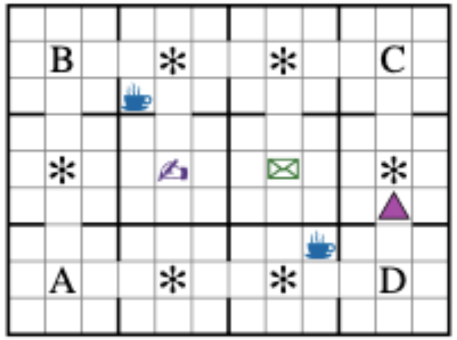
Sample Efficiency



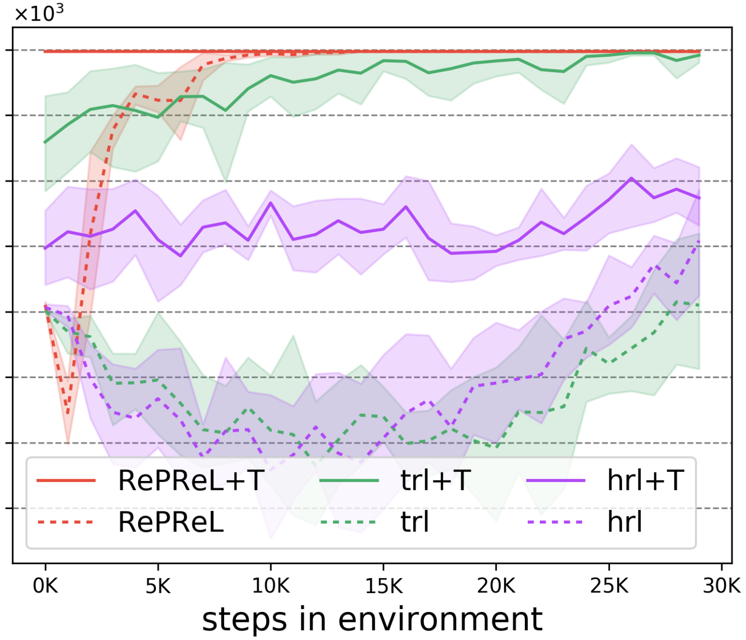




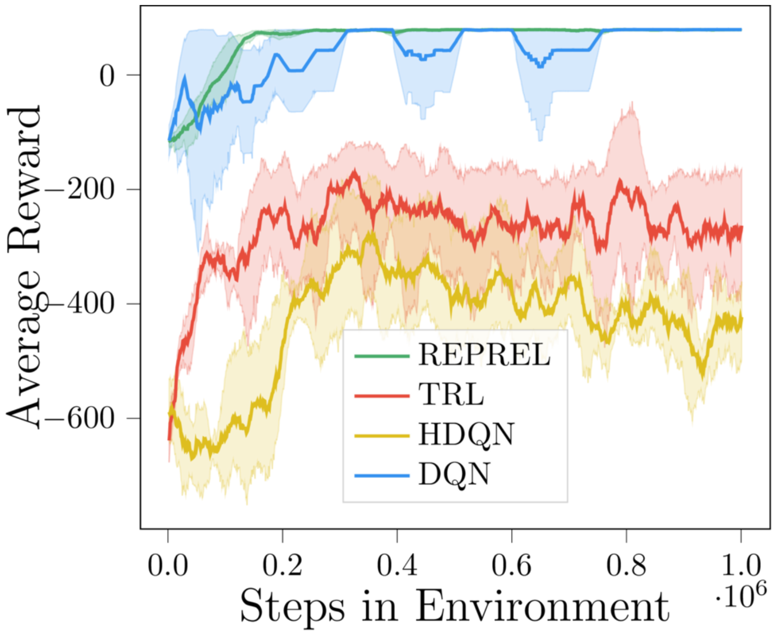
Task Transfer



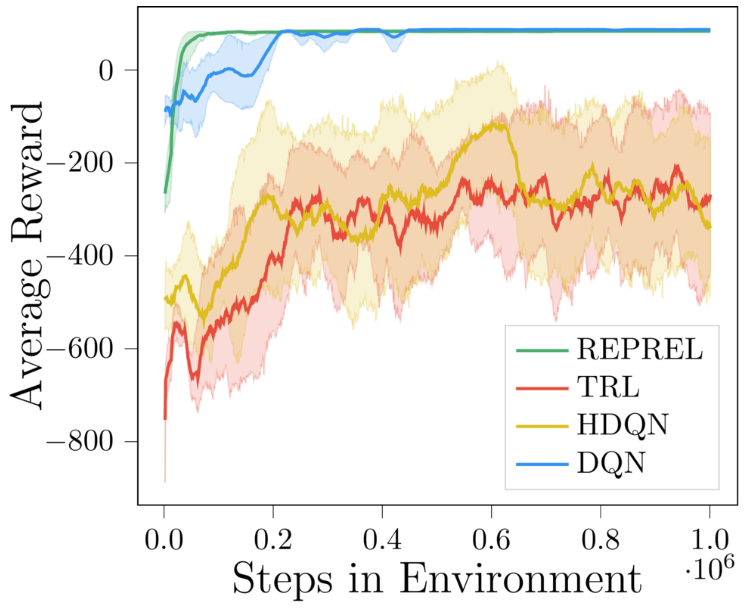
Symbol	Meaning
▲	Agent
*	Furniture
☕	Coffee machine
✉	Mail room
✍	Office
A, B, C, D	Marked locations



Tabular



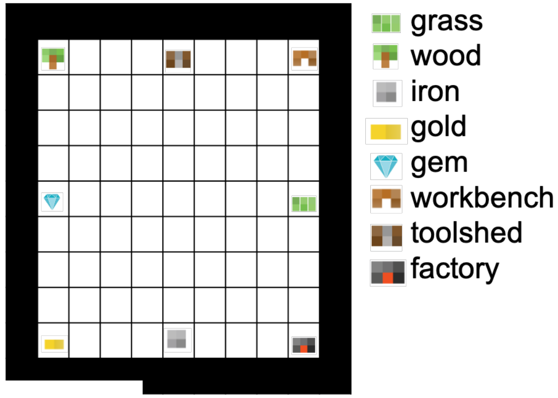
Deliver coffee

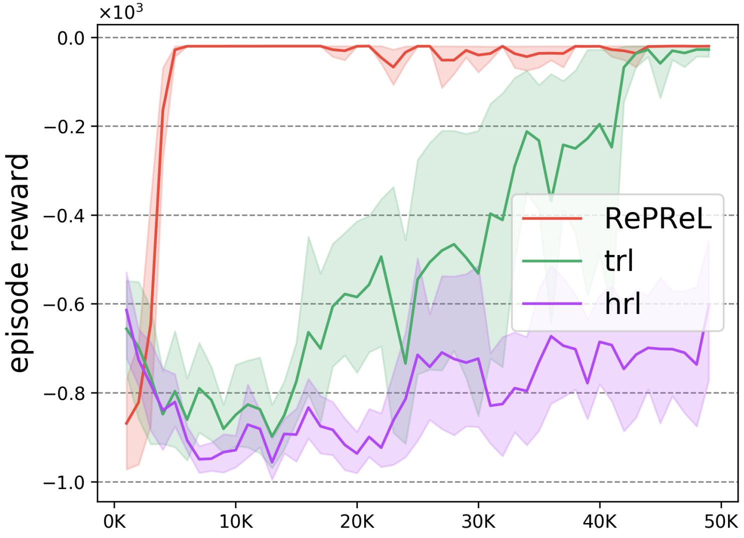


Deliver mail

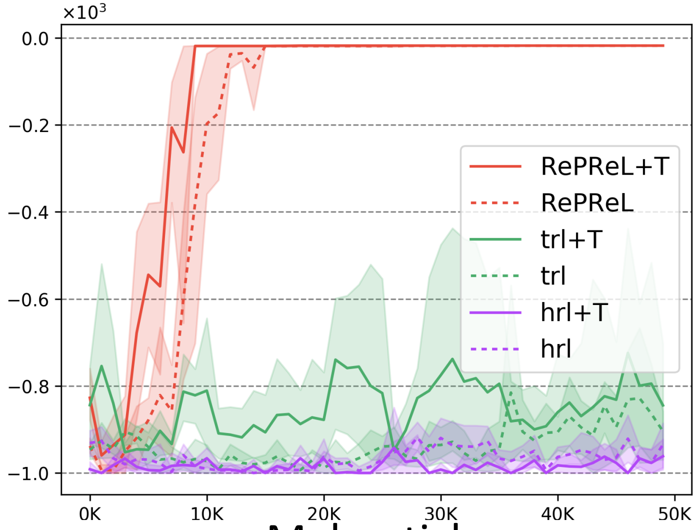
Task Transfer

CRAFT WORLD



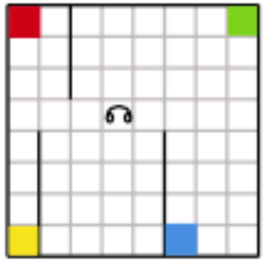


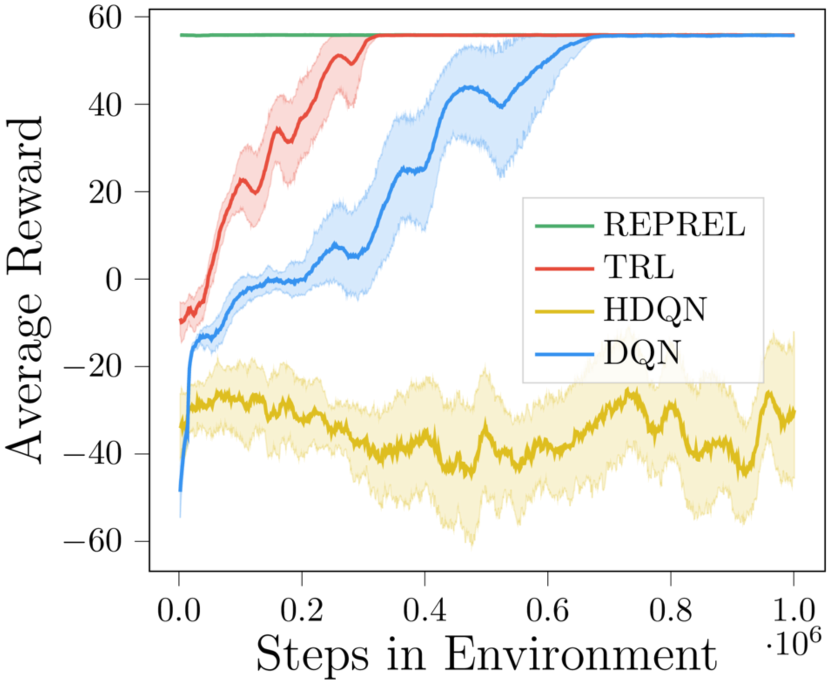
Get wood and iron



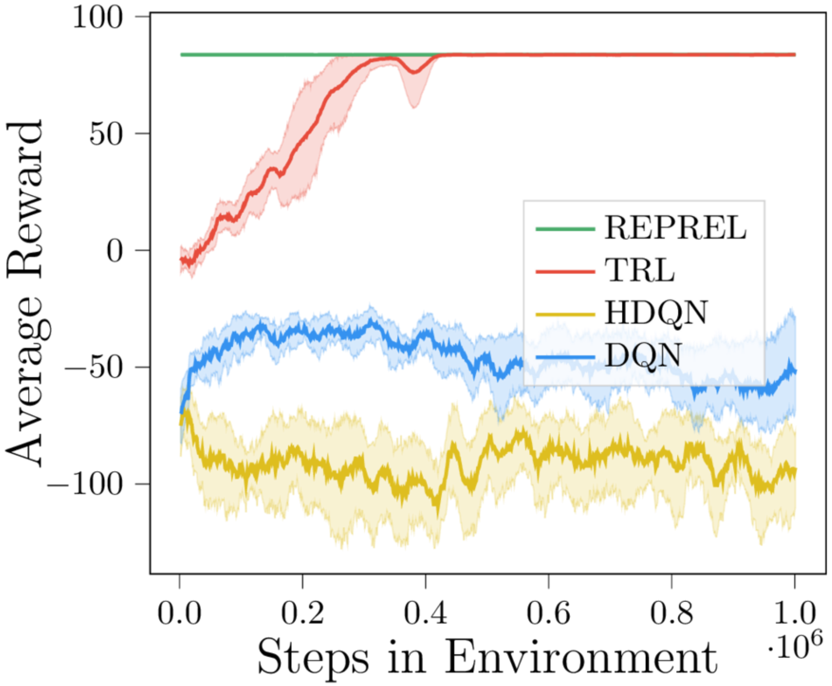
Make stick
(at workbench, needs wood)

Varying # of objects



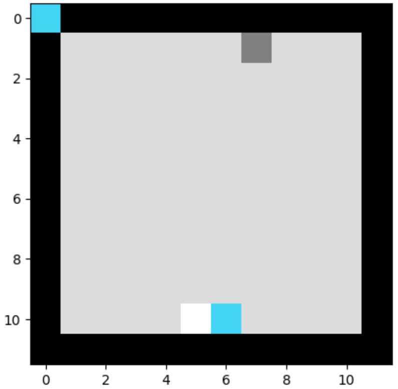


Transport 2 passengers

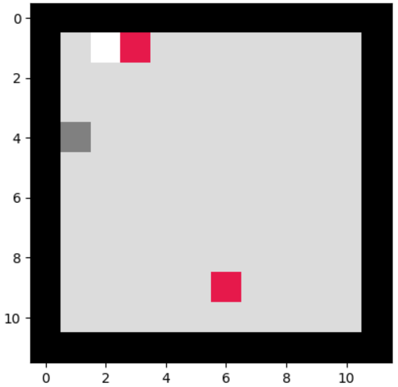


Transport 3 passengers

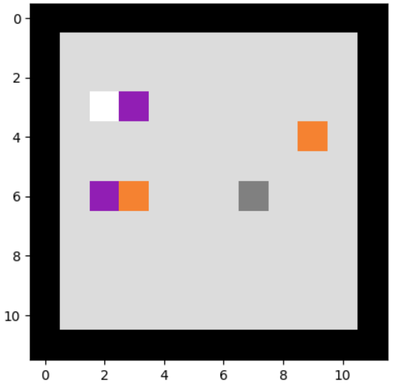
Varying # of objects



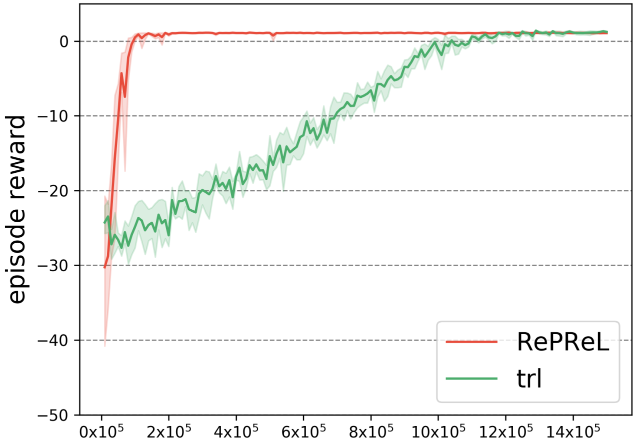
Open lock

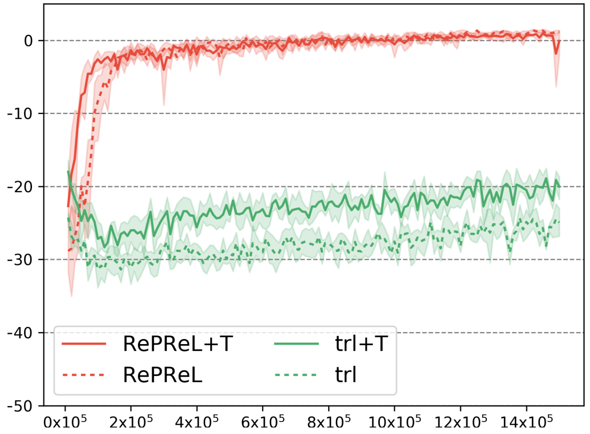


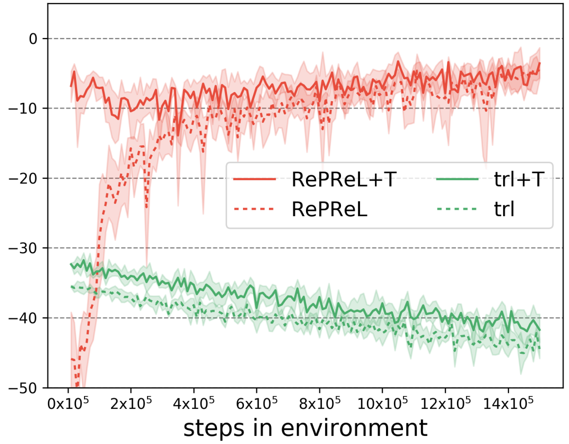
Collect key and open lock



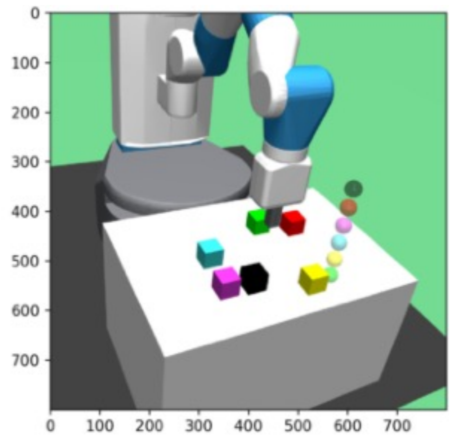
Collect key and
open 2 locks

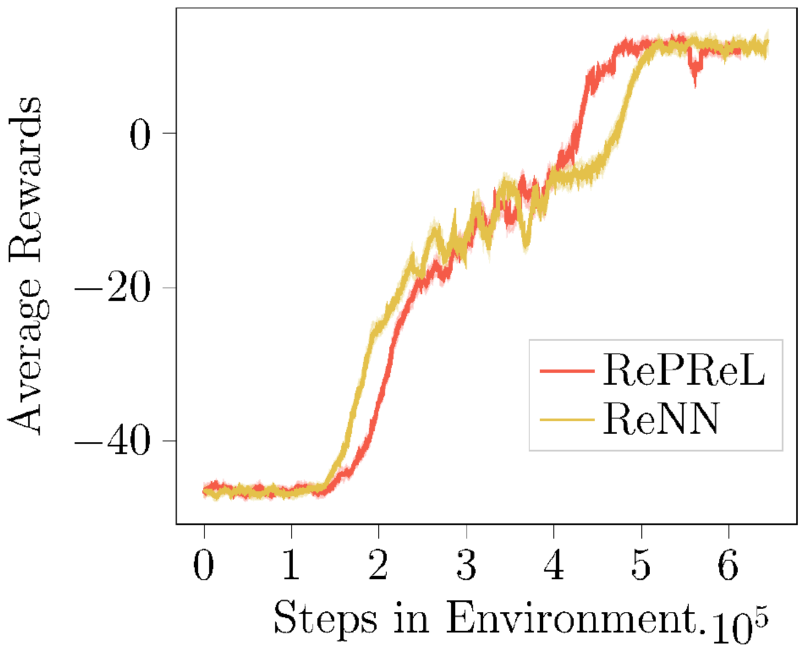




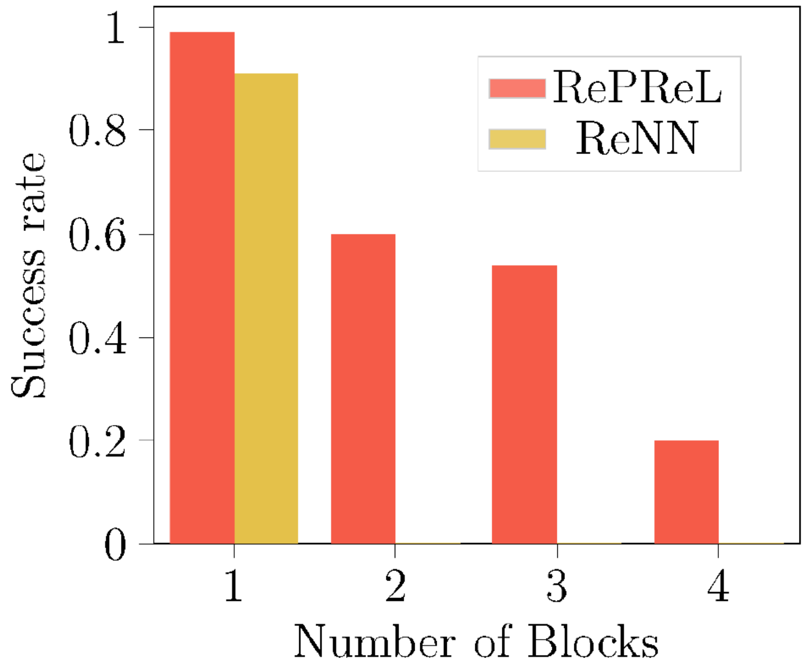


Varying # of objects



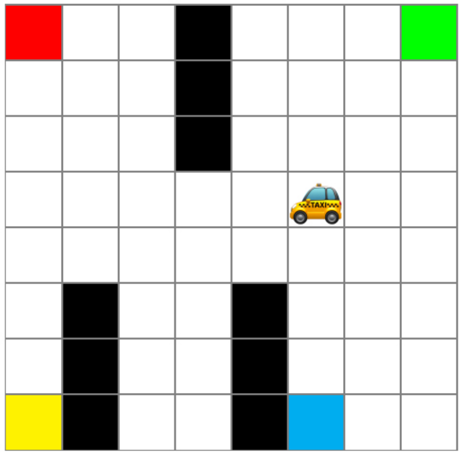


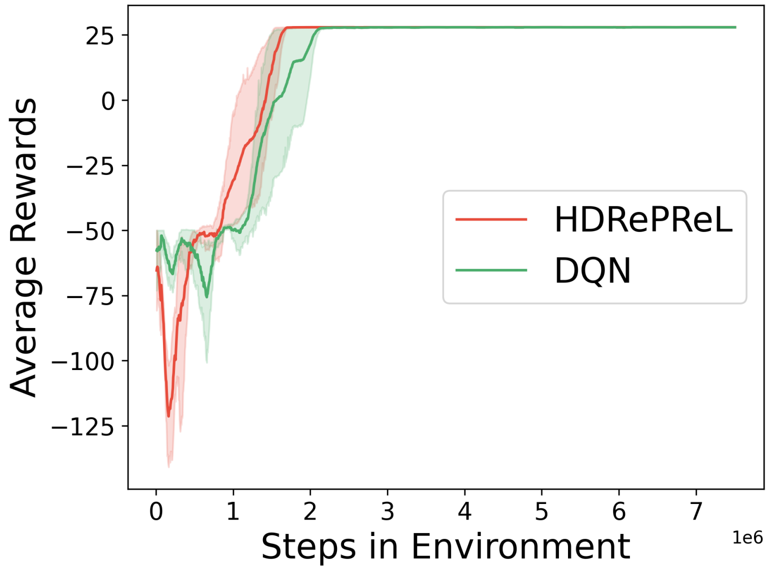
(a)



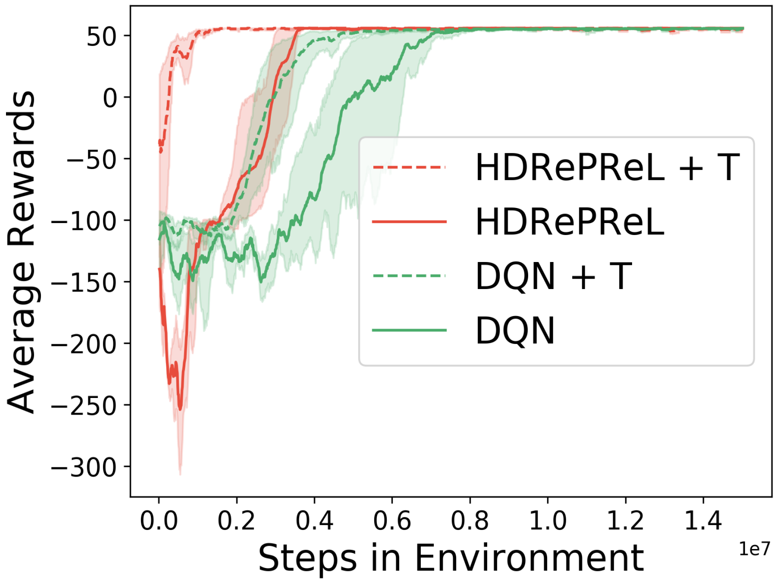
(b)

Multi modal

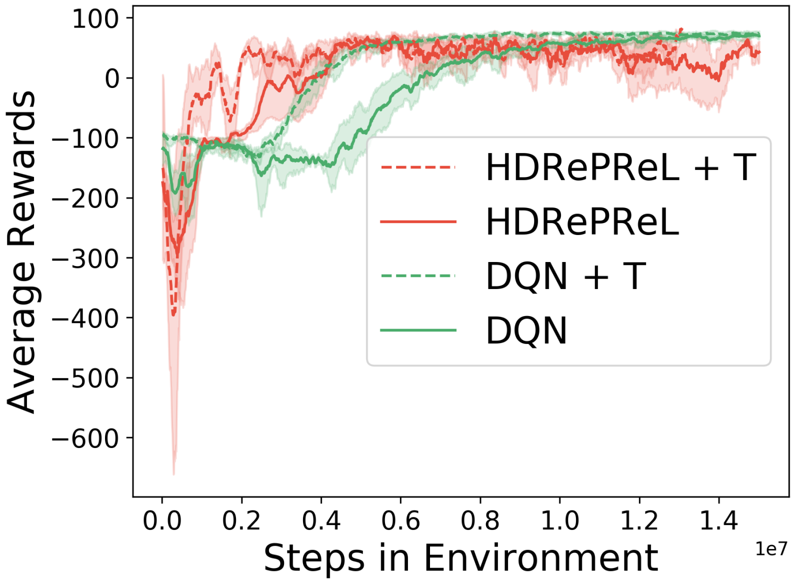




transport one passenger

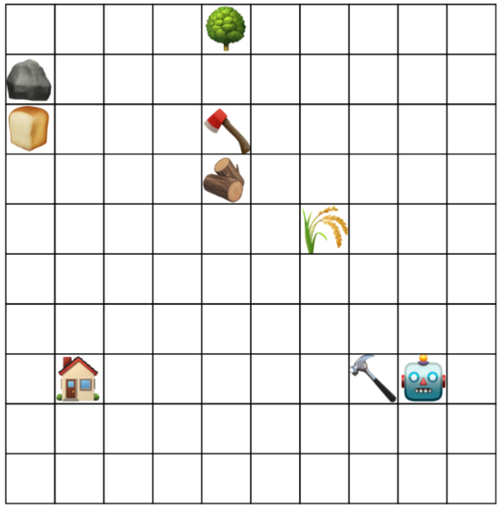


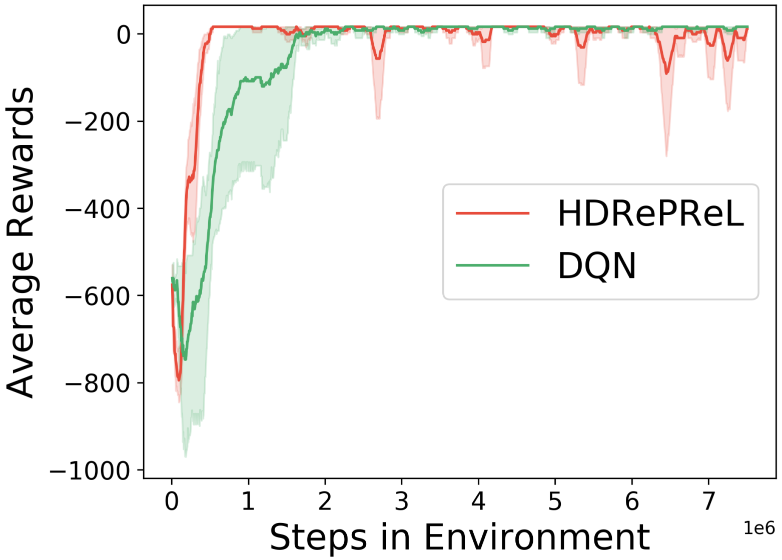
Transport two passengers



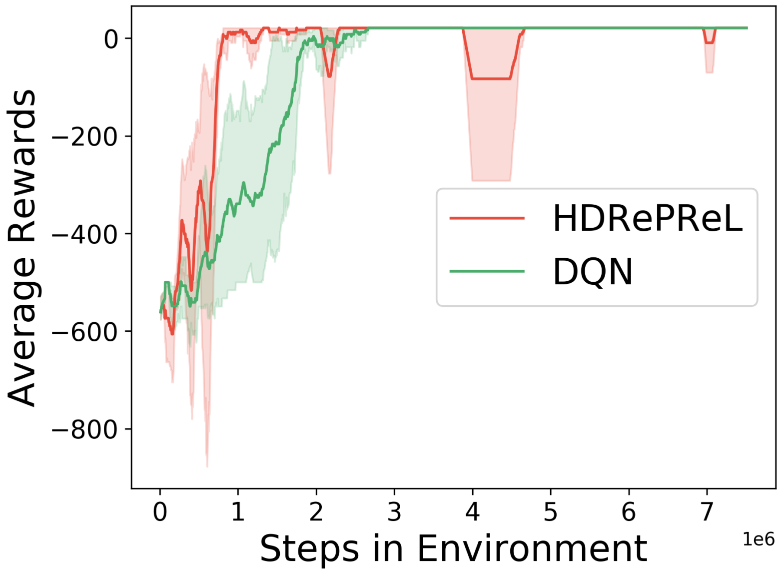
Transport three passengers

Multi modal

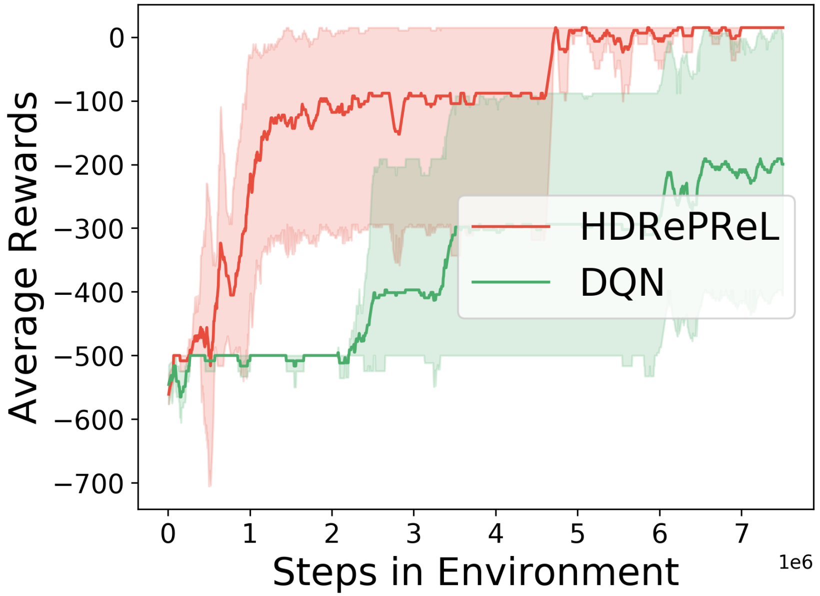




Make Bread



Build a house



Break a rock

Given: Relational sequential decision-making domain

To do: Learn an efficient agent that

- is compositional
- can handle varying # of objects
- can generalize to different tasks
- can support task-specific representations
- can handle multi-modal data

Summary

- Combined a symbolic planner with RL agents
- Provide a batch learning algorithm
- 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
- Most importantly, the framework is planner agnostic and RL algorithm agnostic

Future work

- Refine the D-FOCI statements
- Relax downward refinement
- Partial observability and uncertainty over states
- Boolean task algebra style compositions

Questions?

References

StarAI

Raedt, L.D., Kersting, K., Natarajan, S. and Poole, D., 2016. Statistical relational artificial intelligence: Logic, probability, and computation. *Synthesis lectures on artificial intelligence and machine learning*, *10*(2), pp.1-189.

Raedt, L.D., Dumančić, S., Manhaeve, R., & Marra, G. (2020). From statistical relational to neuro-symbolic artificial intelligence. IJCAI.

Starling lab: Parkinson's Patient

Dhami, D.S., & Soni, A., & Page, D., & Natarajan, S., *Identifying Parkinson's Patients : A Functional Gradient Boosting Approach*, Artificial Intelligence in Medicine (AIME) 2017.

Dhami, D.S., & Das, M., & Natarajan, S., *Beyond Simple Images: Human Knowledge-Guided GANs for Clinical Data Generation*, 8th International Conference on Principles of Knowledge Representation and Reasoning (KR) 2021.

Starling lab: Cohort of pregnant women

Karanam, A., & Hayes, A.L., & Kokel, H., & Haas, D.M., & Radivojac, P., & Natarajan, S., *A Probabilistic Approach to Extract Qualitative Knowledge for Early Prediction of Gestational Diabetes*, 19th International Conference in Artificial Intelligence in Medicine(AIME) 2021.

Pagel, K.A., & Chu, H., & Ramola, R., & Guerrero, R.F., & Chung, J.H., & Parry, S., & Reddy, U.M., & Silver, R.M., & Steller, J.G., & Yee, L.M., & Wapner, R.J., & Hahn, M.W., & Natarajan, S., & Haas, D.M., & Radivojac, P., *Association of Genetic Predisposition and Physical Activity With Risk of Gestational Diabetes in Nulliparous Women*, JAMA Network Open 2022.

Mathur, S. & Karanam, A. & Radivojac, P. & Haas, D.M. & Kersting, K. & Natarajan, S. Exploiting Domain Knowledge as Causal Independencies in Modeling Gestational Diabetes, 28th Pacific Symposium on Biocomputing (PSB) 2023.

Karanam, A., & Mathur, S., & Haas, D.M., & Radivojac, P., & Kersting, K., & Natarajan, S., *Explaining Deep Tractable Probabilistic Models: The sum-product network case*, The Fifth Workshop On Tractable Probabilistic Modeling (TPM) 2022.

References

Starling lab: Drug-Drug Interaction

Odom, P., & Bangera, V., & Khot, T., & Page, D., & Natarajan, S., *Extracting Adverse Drug Events from Text using Human Advice*, Artificial Intelligence in Medicine (AIME) 2015.

Dhami, D.S., & Yan, S., & Kunapuli, G., & Page, D., & Natarajan, S., *Predicting Drug-Drug Interactions from Heterogeneous Data: An Embedding Approach*, 19th International Conference in Artificial Intelligence in Medicine(AIME) 2021

Dhami, D.S., & Kunapuli, G., & Page, D., & Natarajan, S., *Predicting Drug-Drug Interactions from Molecular Structure Images*, AAAI Fall symposium - AI for Social Good 2019.

Starling lab: ChEBI

Das, M., & Ramanan, N., & Doppa, J.R., & Natarajan, S., *Few-Shot Induction of Generalized Logical Concepts via Human Guidance*, Computational Intelligence in Robotics, Frontiers in Robotics and AI 2020.

Starling lab: Social Network

Ramanan, N., & Kunapuli, G., & Khot, T., & Fatemi, B., & Kazemi, S.M., & Poole, D., & Kersting, K., & Natarajan, S., *Structure Learning for Relational Logistic Regression: An Ensemble Approach*, DMKD Journal 2021.

Dhami, D.S., & Yan, S., & Kunapuli, G., & Natarajan, S., *Non-Parametric Learning of Embeddings for Relational Data using Gaifman Locality Theorem*, International Conference on Inductive Logic Programming (ILP) 2021

Kaur, N., & Kunapuli, G., & Natarajan, S., *Non-Parametric Learning of Lifted Restricted Boltzmann Machines*, International Journal of Approximate Reasoning 2020

Starling lab: Games

Kokel, H., & Das, M., & Islam, R., & Bonn, J., & Cai, J., & Dan, S., & Narayan-Chen, A., & Jayannavar, P., & Doppa, J.R., & Hockenmaier, J., & Natarajan, S., & Palmer, M., & Roth, D., *Human-guided Collaborative Problem Solving: A Natural Language based Framework*, In ICAPS. 2021

References

Abstraction

Konidaris, G., 2019. On the necessity of abstraction. *Current opinion in behavioral sciences*, *29*, pp.1-7.

Li, L., Walsh, T.J. and Littman, M.L., 2006, January. Towards a Unified Theory of State Abstraction for MDPs. In *AI&M*.

Planning + RL

Grounds, M. and Kudenko, D., 2005. Combining reinforcement learning with symbolic planning. In AAMAS.

Yang, F., Lyu, D., Liu, B. and Gustafson, S., 2018. Peorl: Integrating symbolic planning and hierarchical reinforcement learning for robust decision-making. arXiv preprint arXiv:1804.07779.

Lyu, D., Yang, F., Liu, B. and Gustafson, S., 2019, July. SDRL: interpretable and data-efficient deep reinforcement learning leveraging symbolic planning. In AAAI.

Jiang, Y., Yang, F., Zhang, S. and Stone, P., 2019, November. Task-motion planning with reinforcement learning for adaptable mobile service robots. In IROS. IEEE.

Eppe, M., Nguyen, P.D. and Wermter, S., 2019. From semantics to execution: Integrating action planning with reinforcement learning for robotic causal problem-solving. Frontiers in Robotics and AI, 6, p.123.

Illanes, L., Yan, X., Icarte, R.T. and McIlraith, S.A., 2020, June. Symbolic plans as high-level instructions for reinforcement learning. In Proceedings of the international conference on automated planning and scheduling (Vol. 30, pp. 540-550).

Lee, J., Katz, M., Agravante, D.J., Liu, M., Klinger, T., Campbell, M., Sohrabi, S. and Tesauro, G., 2022. AI Planning Annotation for Sample Efficient Reinforcement Learning. In PRL@ICAPS 2022.

Goel, S., Shukla, Y., Sarathy, V., Scheutz, M. and Sinapov, J., 2022. RAPid-Learn: A Framework for Learning to Recover for Handling Novelties in Open-World Environments.I n ICDL.

Mitchener, L., Tuckey, D., Crosby, M. and Russo, A., 2022. Detect, Understand, Act: A Neuro-symbolic Hierarchical Reinforcement Learning Framework. Machine Learning, 111(4), pp.1523-1549.

References

Taxi domain

Dietterich, T., 1999. State abstraction in MAXQ hierarchical reinforcement learning. *Advances in Neural Information Processing Systems*, *12*.

FOCI

Natarajan, S., & Tadepalli, P., & Dietterich, T.G., & Fern, A., *Learning First-Order Probabilistic Models with Combining Rules*, Annals of Mathematics and AI, Special Issue on Probabilistic Relational Learning 2008.

Graphical Models

Koller, D. and Friedman, N., 2009. *Probabilistic graphical models: principles and techniques*. MIT press.

Getoor, Lise, and Ben Taskar. "Statistical relational learning." (2007).

Raedt, L.D., Kersting, K., Natarajan, S. and Poole, D., 2016. Statistical relational artificial intelligence: Logic, probability, and computation. *Synthesis lectures on artificial intelligence and machine learning*, *10*(2), pp.1-189.

Neural Predicate

Manhaeve, R., Dumancic, S., Kimmig, A., Demeester, T. and De Raedt, L., 2018. Deepproblog: Neural probabilistic logic programming. *Advances in Neural Information Processing Systems*, *31*.

RePReL and HDRePReL

Kokel, H., & Manoharan, A., & Natarajan, S., & Ravindran, B., & Tadepalli, P., RePReL: Integrating Relational Planning and Reinforcement Learning for Effective Abstraction, In ICAPS 2021a.

Kokel, H., & Manoharan, A., & Natarajan, S., & Ravindran, B., & Tadepalli, P., Deep RePReL-Combining Planning and Deep RL for acting in relational domains, Deep RL Workshop at NeurIPS 2021b.

References

Kokel, H., & Prabhakar, N., & Ravindran, B., & Blasch, E., & Tadepalli, P., & Natarajan, S., Hybrid Deep RePReL: Integrating Relational Planning and Reinforcement Learning for Information Fusion,, In FUSION 2022

Baselines

Sutton, R.S., Precup, D. and Singh, S., 1998, July. Intra-Option Learning about Temporally Abstract Actions. In *ICML*(Vol. 98, pp. 556-564).

Illanes, L., Yan, X., Icarte, R.T. and McIlraith, S.A., 2020, June. Symbolic plans as high-level instructions for reinforcement learning. In ICAPS.

Van Hasselt, H., Guez, A. and Silver, D., 2016, March. Deep reinforcement learning with double q-learning. In AAAI.

Schulman, J., Wolski, F., Dhariwal, P., Radford, A. and Klimov, O., 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*.

Li, R., Jabri, A., Darrell, T. and Agrawal, P., 2020, May. Towards practical multi-object manipulation using relational reinforcement learning. In *ICRA*.

Andrychowicz, M., Wolski, F., Ray, A., Schneider, J., Fong, R., Welinder, P., McGrew, B., Tobin, J., Pieter Abbeel, O. and Zaremba, W., 2017. Hindsight experience replay. *NeurIPS*,

Haarnoja, T., Zhou, A., Abbeel, P. and Levine, S., 2018, July. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *ICML* PMLR.

Capacity loss in Deep RL

Lyle, C., Rowland, M. and Dabney, W., 2022. Understanding and preventing capacity loss in reinforcement learning. *ICLR*.

THANKS

starling.utdallas.edu
harshakokel.com

StaRLinG LAB

@starling_lab
@harsha_kokel