Unsupervised Subgoal Discovery Method for Learning Hierarchical Representations

Jacob Rafati

Ph.D., Electrical Engineering and Computer Science (EECS)
Computational Cognitive Neuroscience Laboratory (CCNL)
http://rafati.net

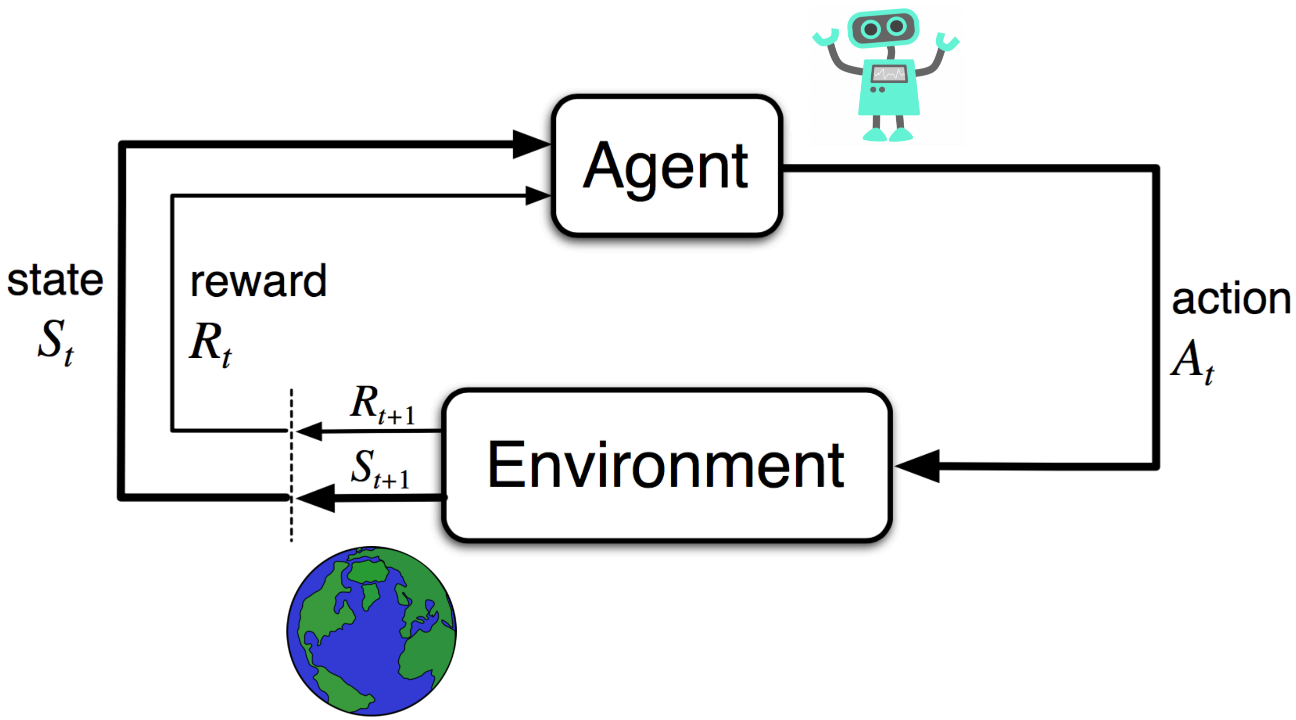
Co-authored with **David C. Noelle**

Professor and Chair of Cognitive Information and Sciences
Founding Faculty of EECS & CSE, Director of CCNL

University of California, Merced

Workshop on Structure and Priors in Reinforcement Learning (SPiRL 2019)
7th International Conference on Learning Representations (ICLR 2019)

Reinforcement Learning



Reinforcement learning (RL) is learning how to map situations (**states**) to **agent**'s decisions (**actions**) to **maximize future rewards (return)** by **interaction** with an **unknown environment.**

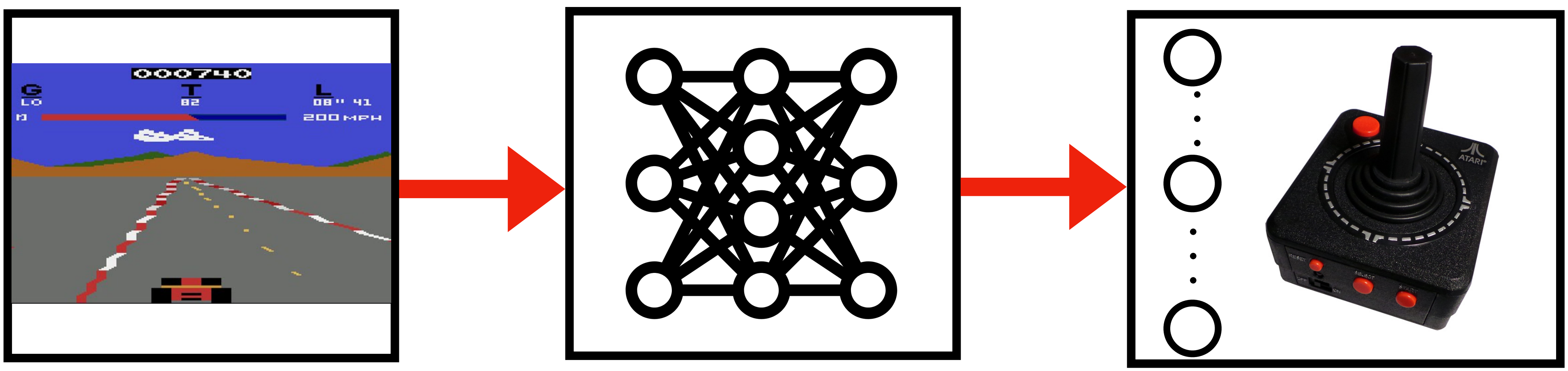
Experience (*s, a, r, s'*) as Data.

Sutton and Barto (2017). Reinforcement Learning: An Introduction. MIT Press, Cambridge, MA, USA, 2nd edition.

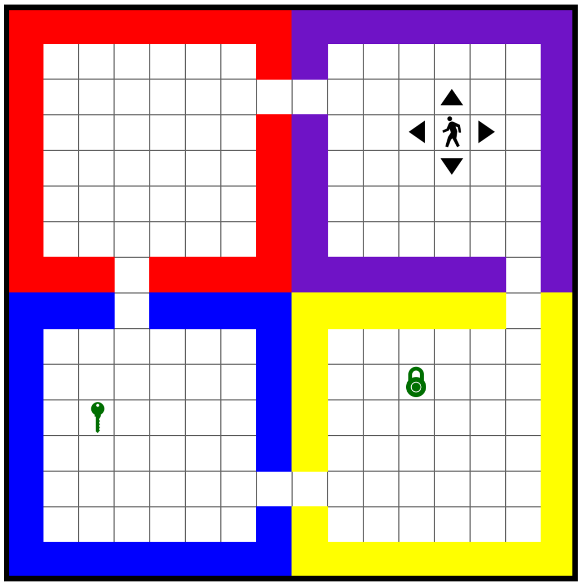
Generalization

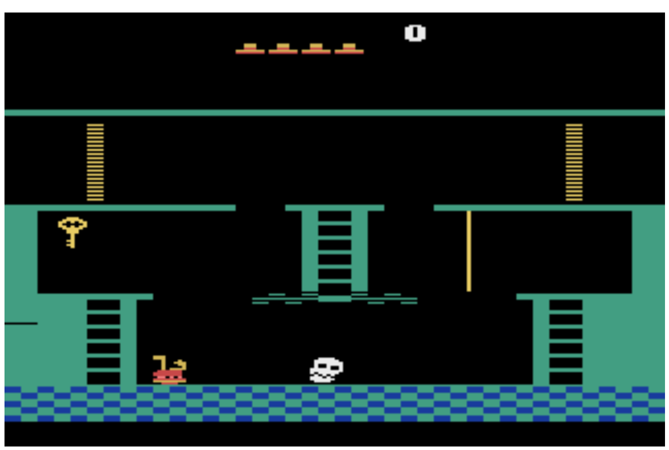
State → **Parameterized Function Approximator** → **Value Function Expectation of Return (Game Scores)**

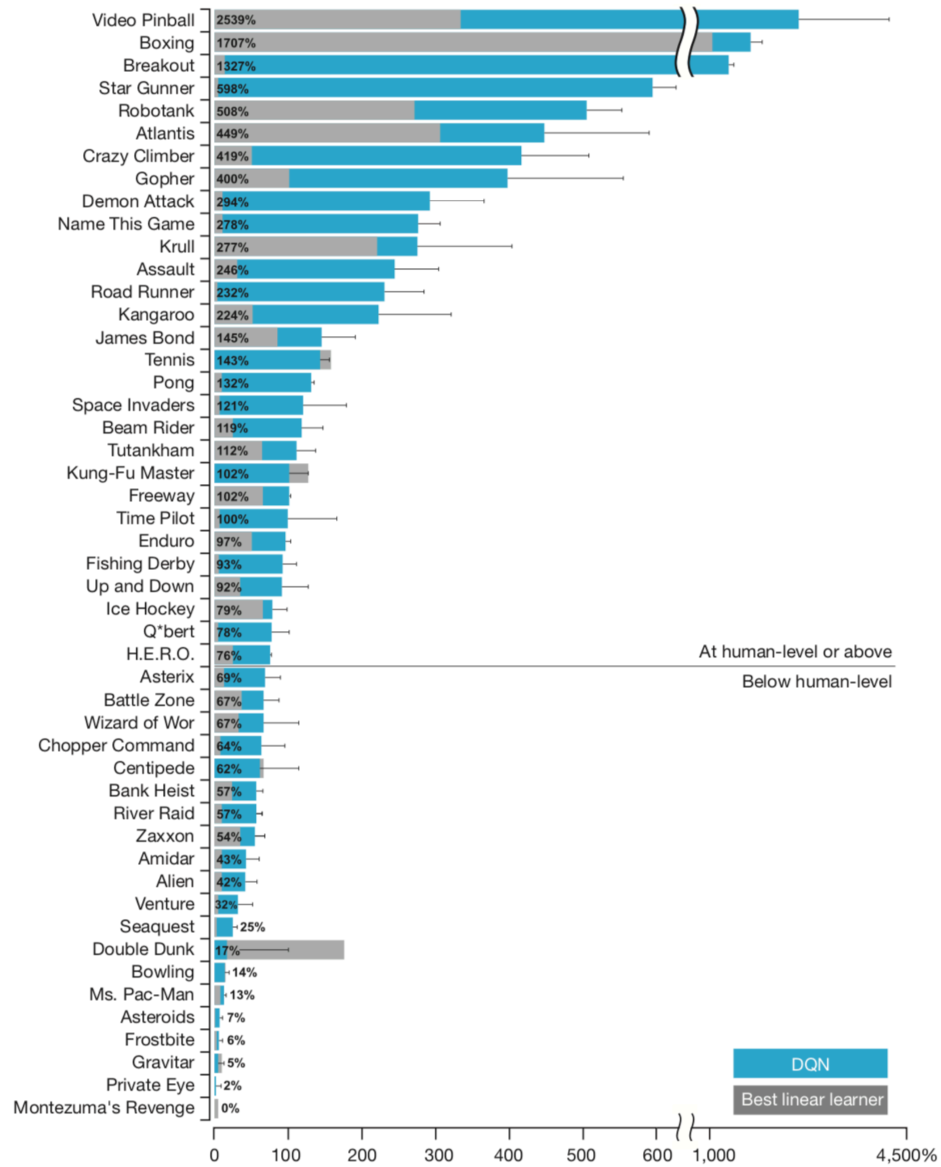
s $\quad$ w $\quad$ $q(s, a_i; w)$



Success in easy tasks, Failure in more complex task







Mnih, et al. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540):529–533.

Learning Representations in model-free HRL

- **Temporal Abstraction**

Learning to operate over different levels of *temporal abstraction*. Learning a meta-policy to choose a proper subgoal.

- **Intrinsic Motivation Learning**

Efficiently exploring the state-space while learning reusable *subpolicies* (skills) through the *intrinsic motivation learning*. The intrinsic critic sends intrinsic rewards based on attaining subgoals.

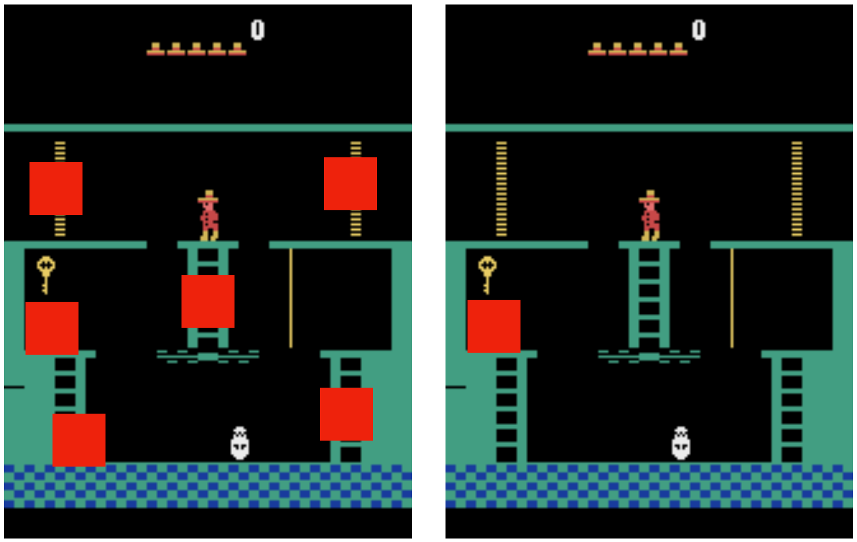
- **Automatic Subgoal Discovery**

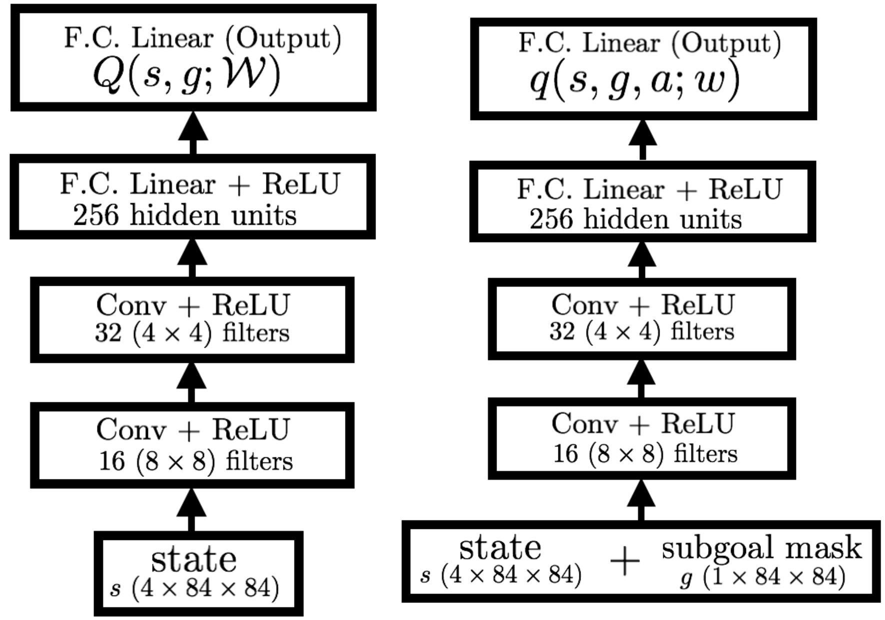
Automatic Subgoal Discovery in large-scale tasks with sparse delayed feedback within model-free HRL framework.

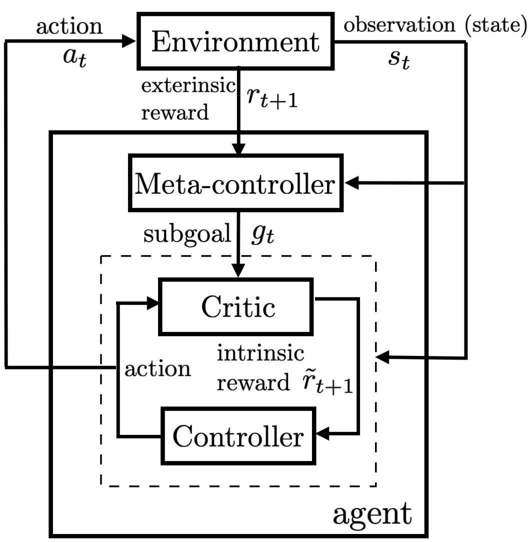
- **Learning hierarchical representation of model-free HRL in a unified approach**

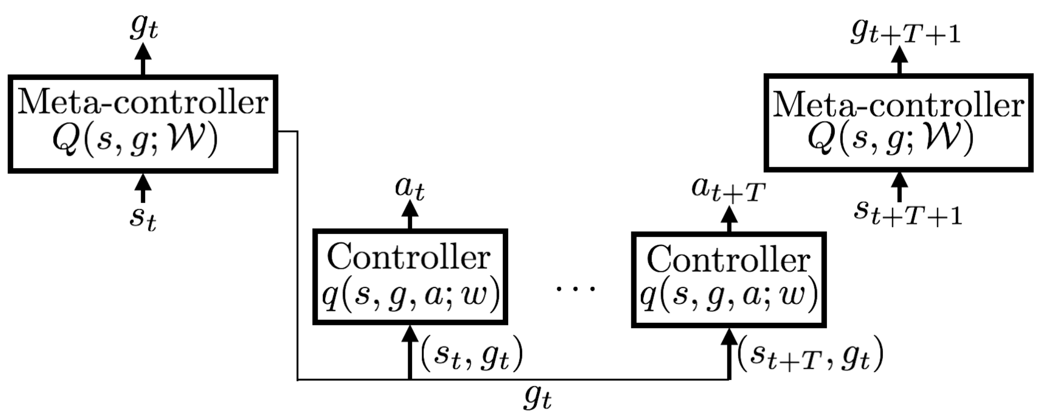
Integration of temporal abstraction, intrinsic motivation learning and subgoal discovery in one unified algorithm.

Meta-controller/Controller Framework









Kulkarni et al. (2016). Hierarchical deep reinforcement learning: Integrating temporal abstraction and intrinsic motivation. NeurIPS.

Unsupervised Subgoal Discovery

Properties:

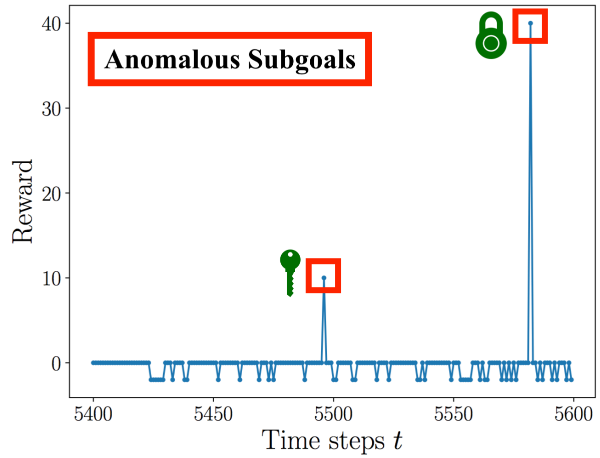
- It is close to a rewarding state.
- It represents a set of states, at least some of which tend to be along a state transition path to a rewarding state.

Hypothesis: We can use unsupervised learning methods to find useful subgoals based on a memory of the agent's experiences (rewards and visited states).

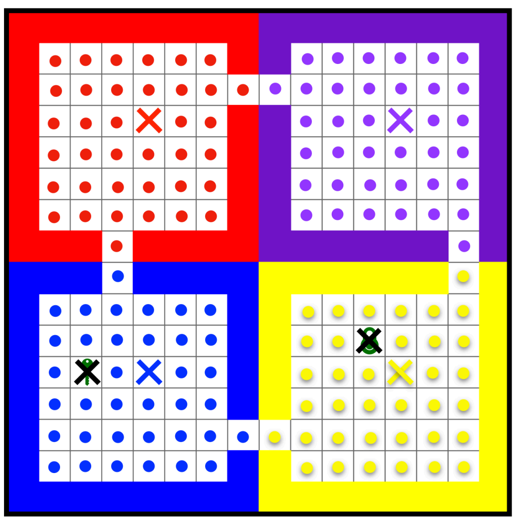
  - Centroids of K-means clusters (e.g. rooms)
  - Outliers as potential subgoals (e.g. key, box)
  - Boundary of two clusters (e.g. doorway)

Unsupervised Subgoal Discovery

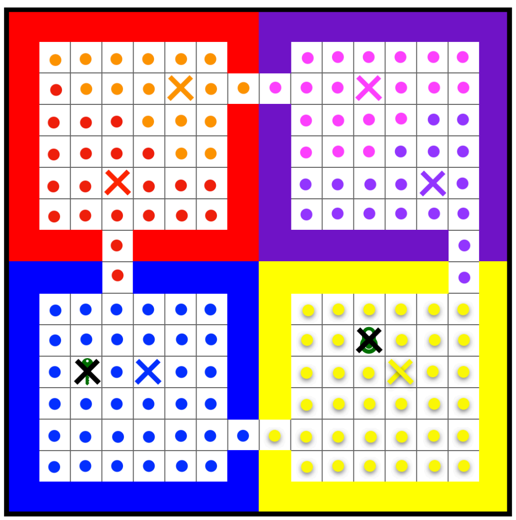
Anomaly Detection



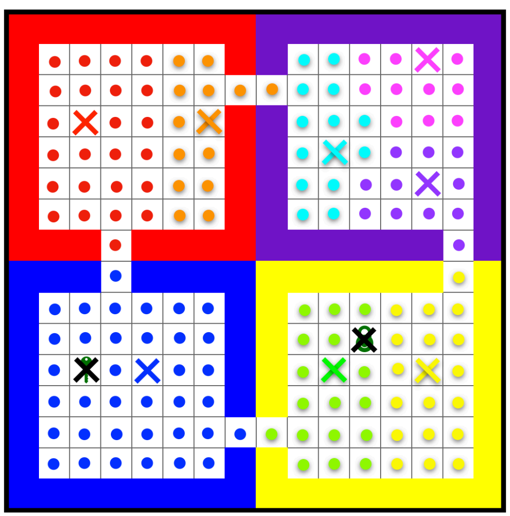
K-Means Clustering



$K = 4$

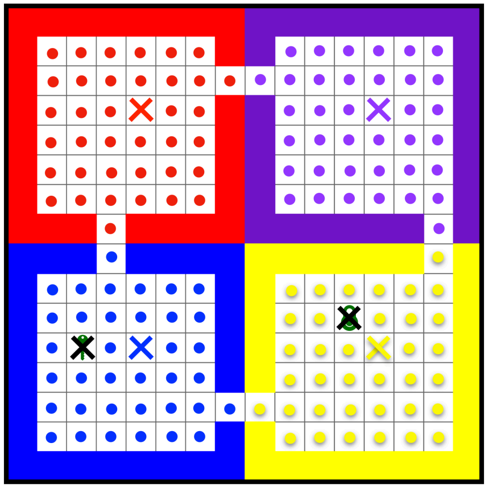


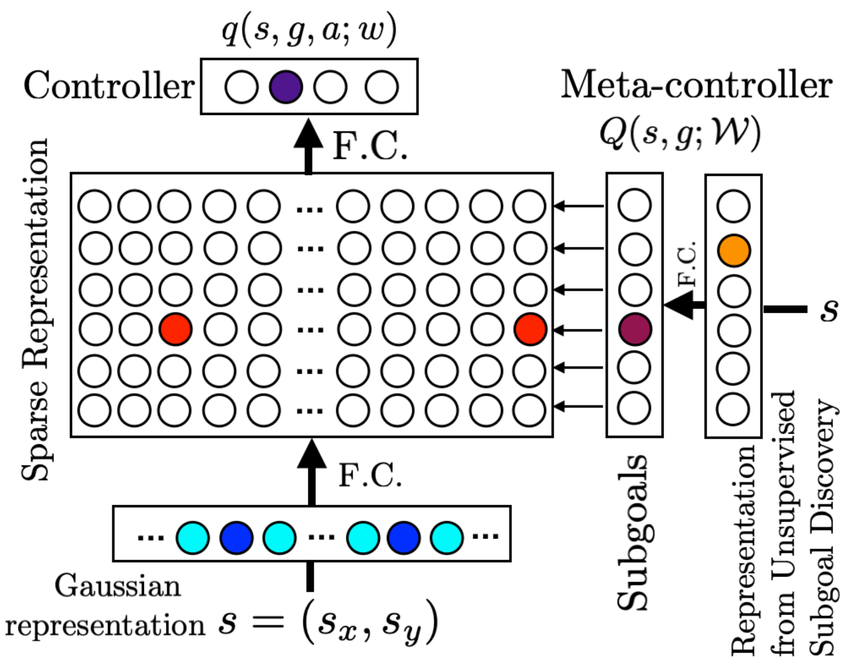
$K = 6$

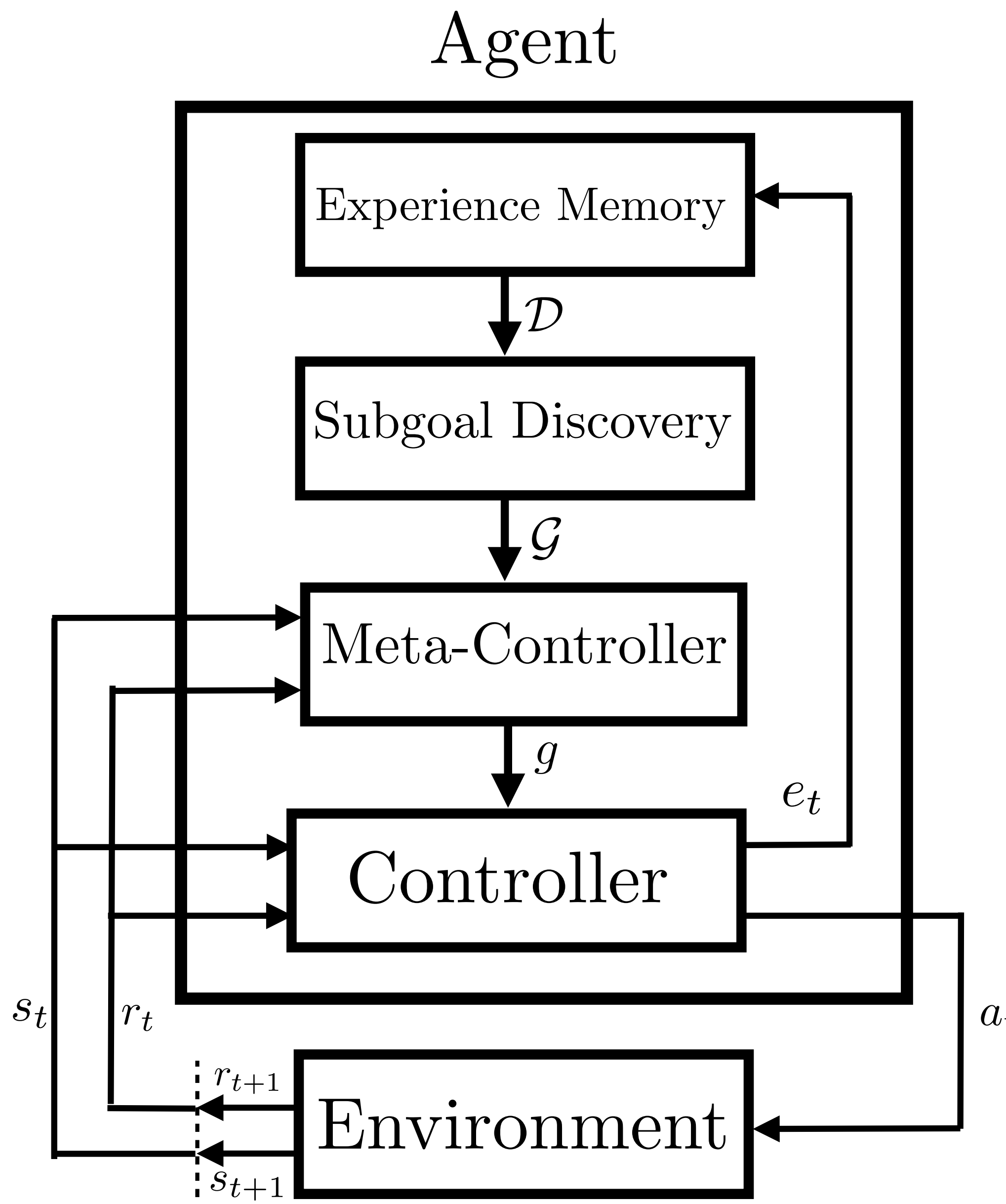


$K = 8$

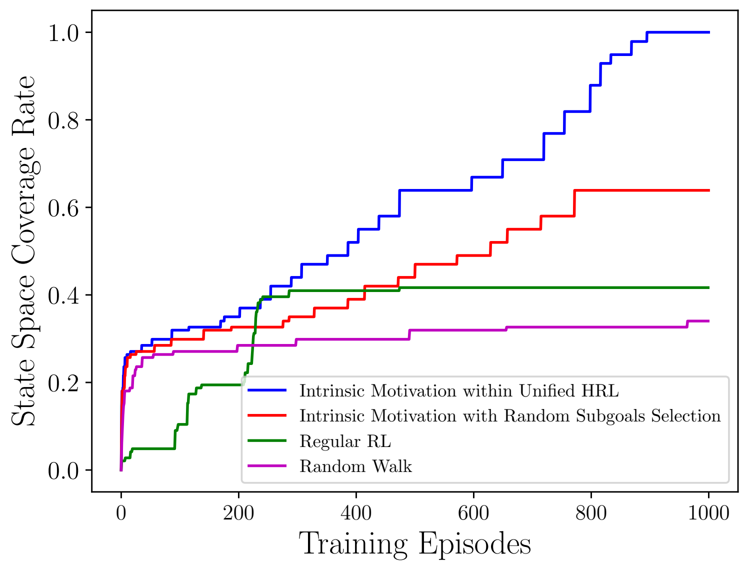
Unified Model-Free HRL

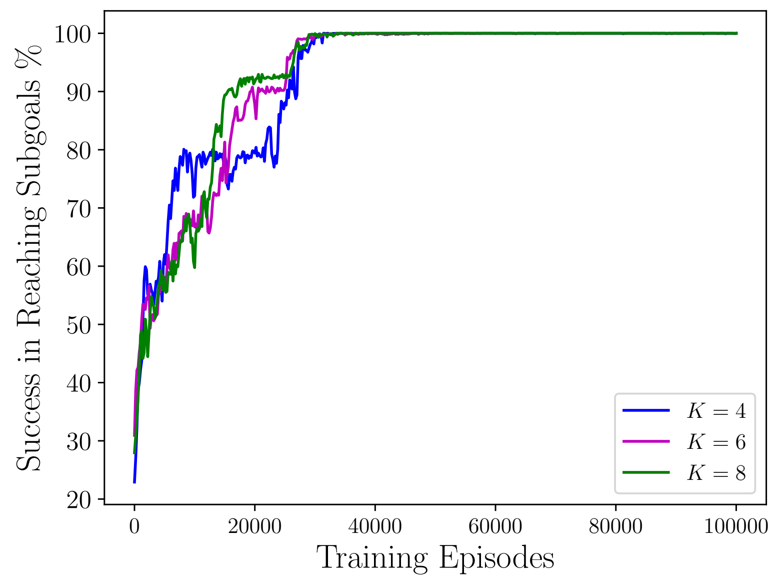


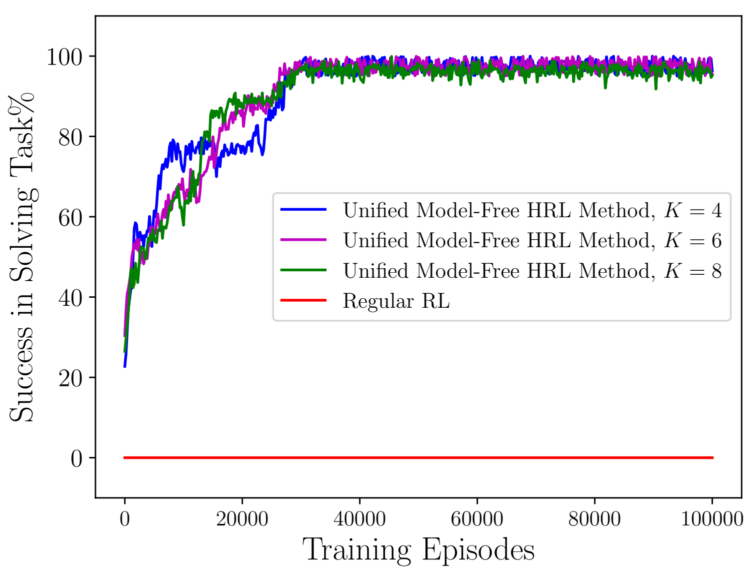


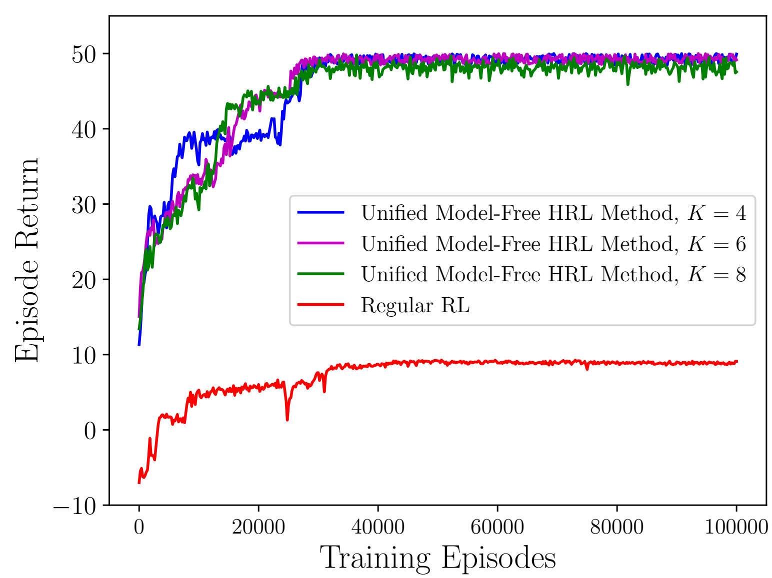


Results — 4-Rooms task





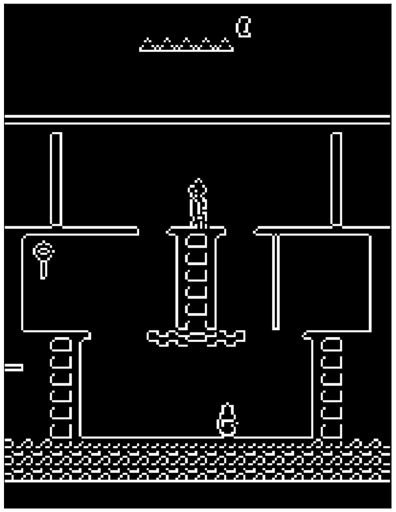




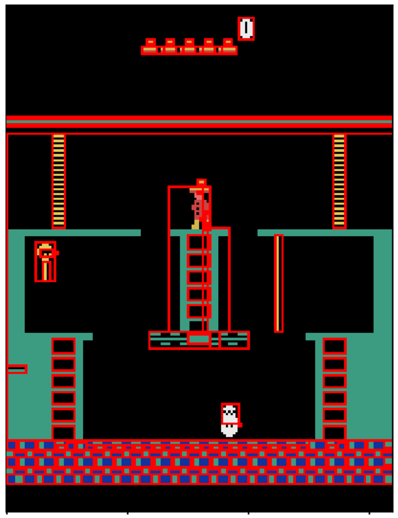
Montezuma's Revenge

Initial Subgoals

Edge Detection

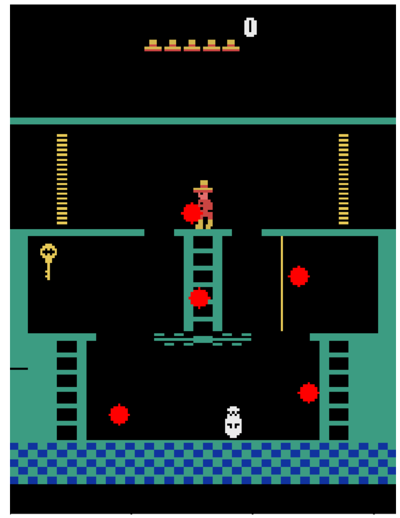


Bounding Box

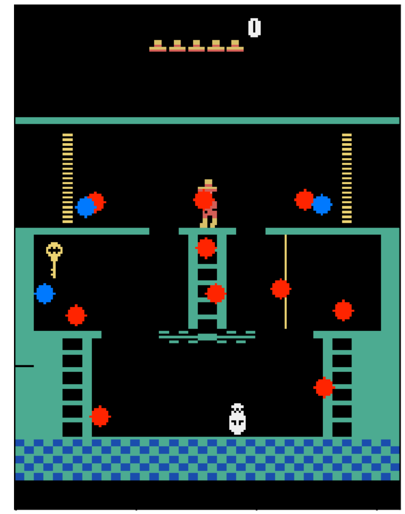


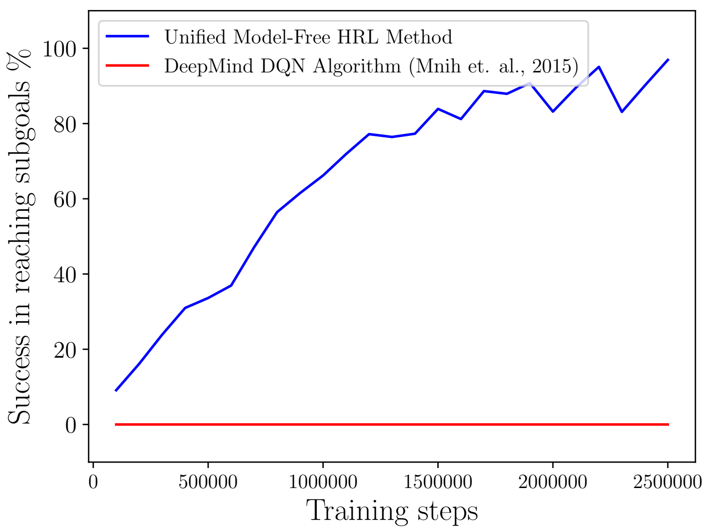
Unsupervised Subgoal Discovery

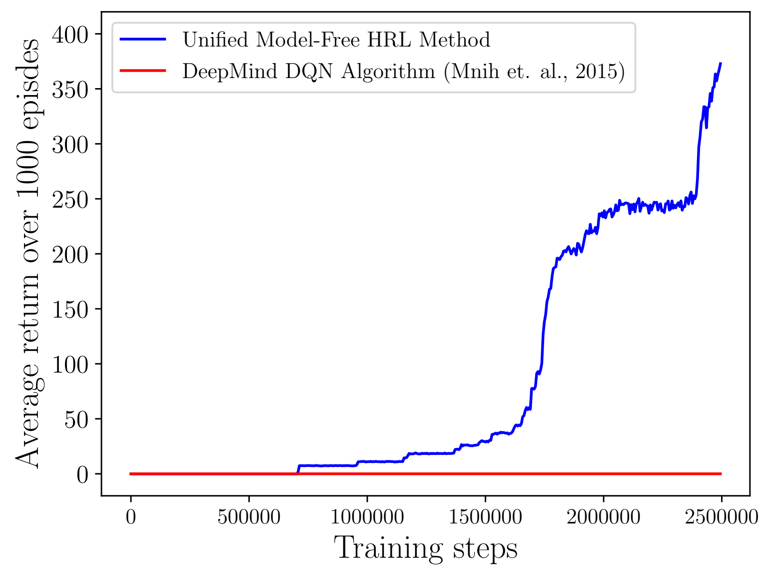
Random Walk



Our Method

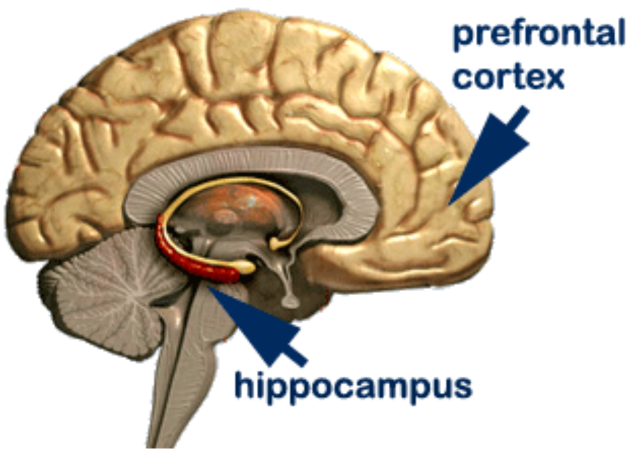


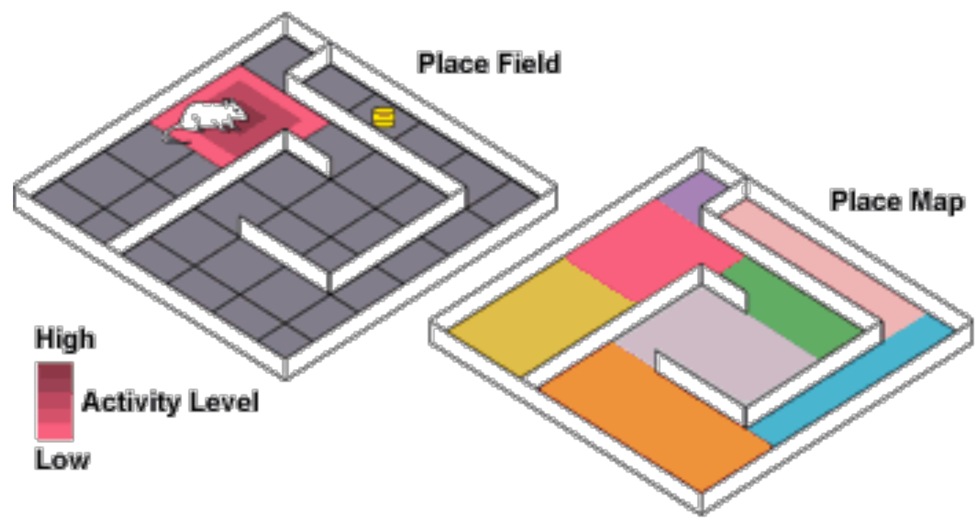




Neural Correlates of Unsupervised Subgoal Discovery

- Temporal abstraction in HRL might map onto regions within the dorsolateral and orbital prefrontal cortex (PFC).
- More recent discoveries reveal a potential role for medial temporal lobe structures, including the hippocampus, in planning and spatial navigation, utilizing a hierarchical representation of space.
- There are evidences that hippocampus serve in model-based and model-free HRL with both flexibility and computational efficiency.
- Place cells in the dorsal hippocampus represent small regions while those in the ventral hippocampus represent larger regions.





Strange et al. (2014). Functional organization of the hippocampal longitudinal axis. Nature Reviews Neuroscience, 15(10):655–669.
Chalmers et al. (2016). Computational properties of the hippocampus increase the efficiency of goal-directed foraging through hierarchical reinforcement learning. Frontiers in Computational Neuroscience, 10.
Botvinick et al. (2009). Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. Cognition, 113(3).
Botvinick, M. and Weinstein, A. (2014). Model-based hierarchical reinforcement learning and human action control. Philosophical Transactions of the Royal Society B: Biological Sciences, 369.

Conclusions

- We proposed and demonstrated a novel model-free method for subgoal discovery using unsupervised learning over a small memory of experiences (trajectories) of the agent.
- When combined with an intrinsic motivation learning mechanism, this method learns subgoals and skills together, based on experiences in the environment.
- Intrinsic motivation learning provides efficient exploration scheme in tasks with sparse rewards that leads to successful subgoal discovery.
- We offered a unified approach for learning hierarchical representations in a model-free HRL framework. This method is scalable to larger scale problems.

Publications

- Jacob Rafati, David C. Noelle. (2019). Unsupervised Subgoal Discovery Method for Learning Hierarchical Representations. In 7th International Conference on Learning Representations, ICLR 2019 Workshop on "Structure & Priors in Reinforcement Learning", New Orleans, LA, USA.
- Jacob Rafati, David C. Noelle. (2019). Unsupervised Methods For Subgoal Discovery During Intrinsic Motivation in Model-Free Hierarchical Reinforcement Learning. In 33rd AAAI Conference on Artificial Intelligence (AAAI-19). Workshop on Knowledge Extraction From Games. Honolulu, Hawaii. USA.
- Jacob Rafati, and David C. Noelle (2019). Learning Representations in Model-Free Hierarchical Reinforcement Learning. In 33rd AAAI Conference on Artificial Intelligence (AAAI-19), Honolulu, Hawaii.
- Jacob Rafati, and David C. Noelle (2019). Learning Representations in Model-Free Hierarchical Reinforcement Learning. arXiv e-print (arXiv:1810.10096).

Questions and Feedbacks

For paper, code, slides:

`http://rafati.net`

Email:

`yrafati@gmail.com`