Towards Jumpy Planning

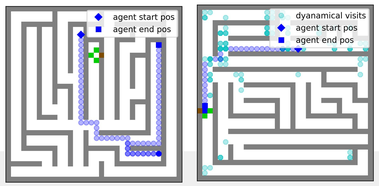
Akilesh B [1,2] **Suriya Singh** [2,3] **Anirudh Goyal** [1,2] **Alexander Neitz** [4] **Aaron Courville** [1,2]

[1] Universite de Montreal [2] Mila [3] Polytechnique Montreal [4] MPI for Intelligent Systems

1. Overview

- Model-free RL: high sample inefficiency and ignorance of the environment dynamics.
- Model-based RL at the scale of time-steps: compounding errors and high computational requirements.

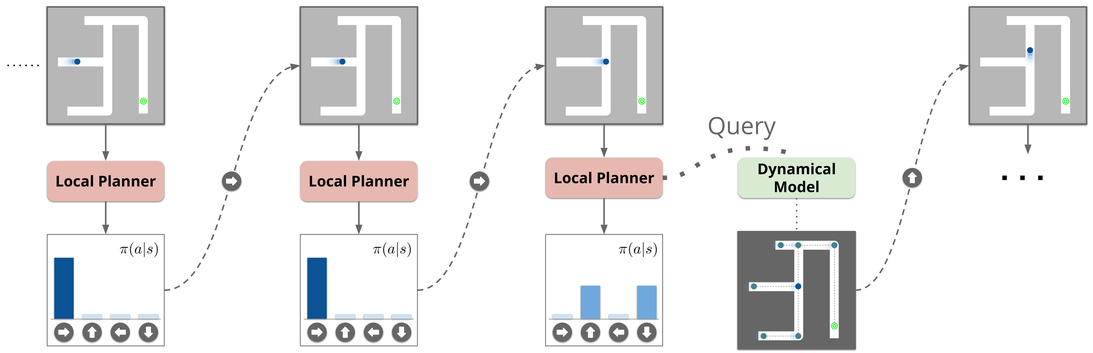
Hierarchical Reinforcement Learning framework [1, 2] address limitations in classic RL through sub-tasks and abstract actions.



This Work

Use a model-based planner together with a goal-conditioned policy trained with model-free learning. We use a model-based planner that operates at higher levels of abstraction i.e., *decision states* and use model-free RL between the decision states.

2. Jumpy Planning

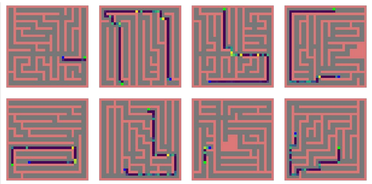


Decision States [3] (*aka* subgoal)

states where the agent's policy has high entropy

$$-\sum_{a\in\mathcal{A}} \pi(a|s) \ln \pi(a|s) > \tau$$

We fix τ such that a tiny fraction of states are chosen as decision



Dynamical Models

$\mathrm{M}\colon (\mathrm{s}, \mathrm{a}) \rightarrow s'$

$\rightarrow$ **argmax** or **sample** successively query M in BFS fashion until the goal state is encountered or maximum search depth is reached.

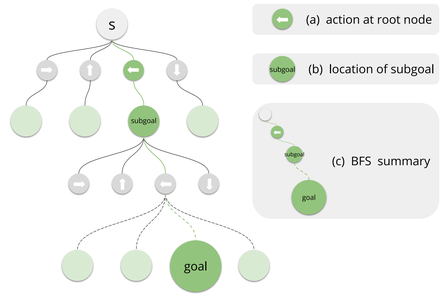
Jumpy Dataset (s, a, s')

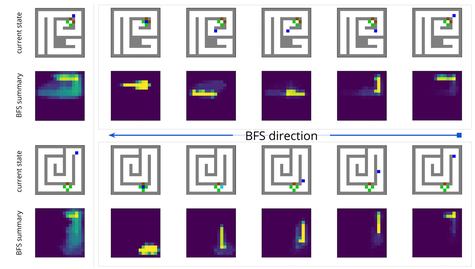
$\sum_{a\in\mathcal{A}} \pi(a|s') \ln \pi(a|s') > \tau$ or $\Delta_{min}T \leq dist(s, s') \leq \Delta_{max}T$

Funnel all intermediate states leading to the same s'

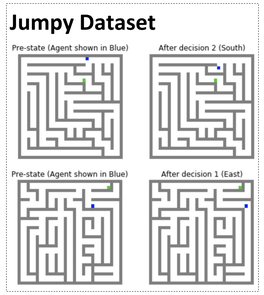
3. Jumpy Planning with Dynamical Models

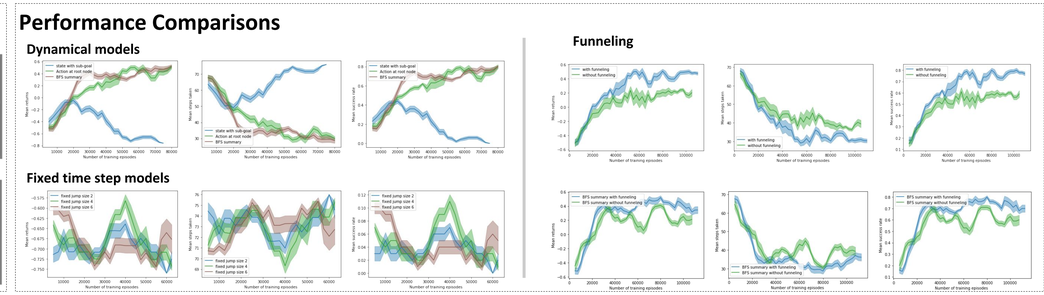
The result of query is further passed to the agent to take action at current decision state

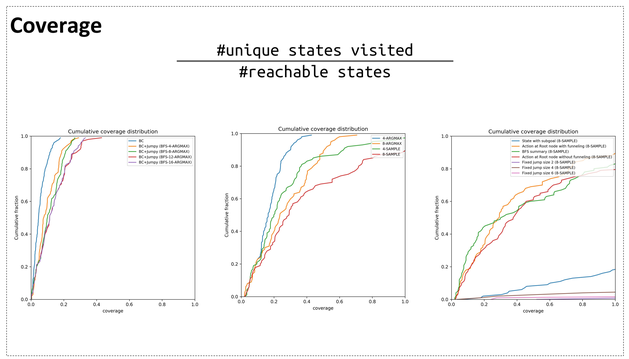


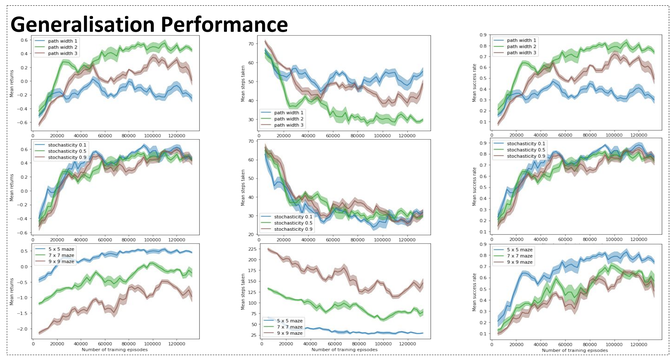


4. Results









References

1) Sutton et. al. Between mdps and semi-mdps: A framework for temporal abstraction in reinforcement learning. *AI Journal*.
2) Hoang Le et al. Hierarchical Imitation and Reinforcement Learning. *ICML 2018.*
3) Goyal et al. InfoBot, Transfer and Exploration via the Information Bottleneck. *ICLR 2019*.

email: akilesh.badrinaaraayanan@umontreal.ca

ICML | 2019
MBRL Workshop