Learning to Navigate in Mazes with Novel Layouts using Abstract Top-down Maps

Linfeng Zhao [1] Lawson L.S. Wong [1]

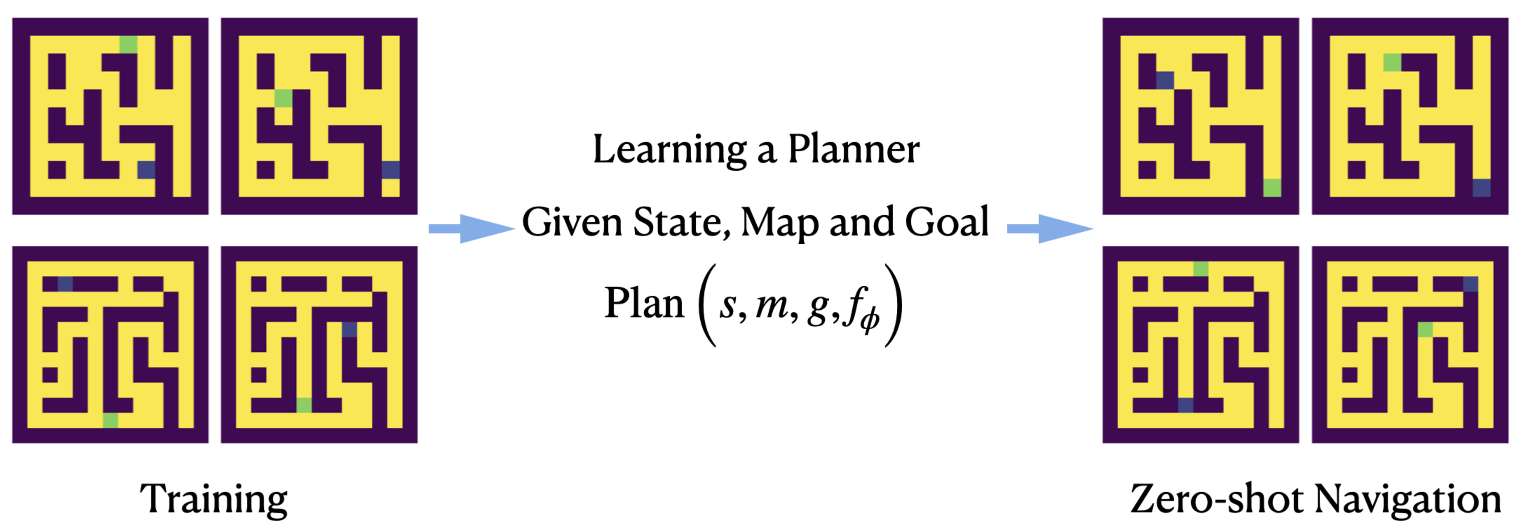
[1]Khoury College of Computer Sciences, Northeastern University

Overview

- We study maze navigation using abstract 2-D maps, inspired by how humans use paper maps to navigate unfamiliar cities.
- Our goal is to achieve zero-shot generalization on environments with unseen layouts, enabling navigation in novel maps and goals without prior exploration.
- We propose a model-based approach that learns a task-conditioned hypermodel end-to-end, utilizing Monte Carlo Tree Search (MCTS) for planning.
- Our method improves robustness to noise and supports hierarchical navigation, allowing effective generalization to long-distance goals on unseen maps

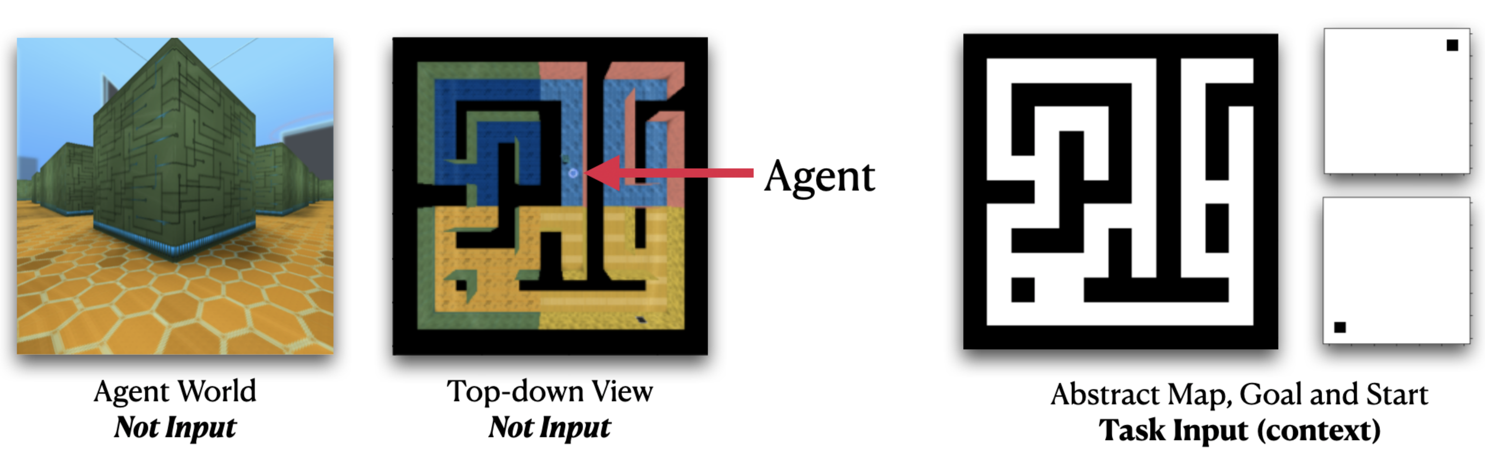
Problem Statement

Setup: Learning Map-based *Maze* Navigation Behavior



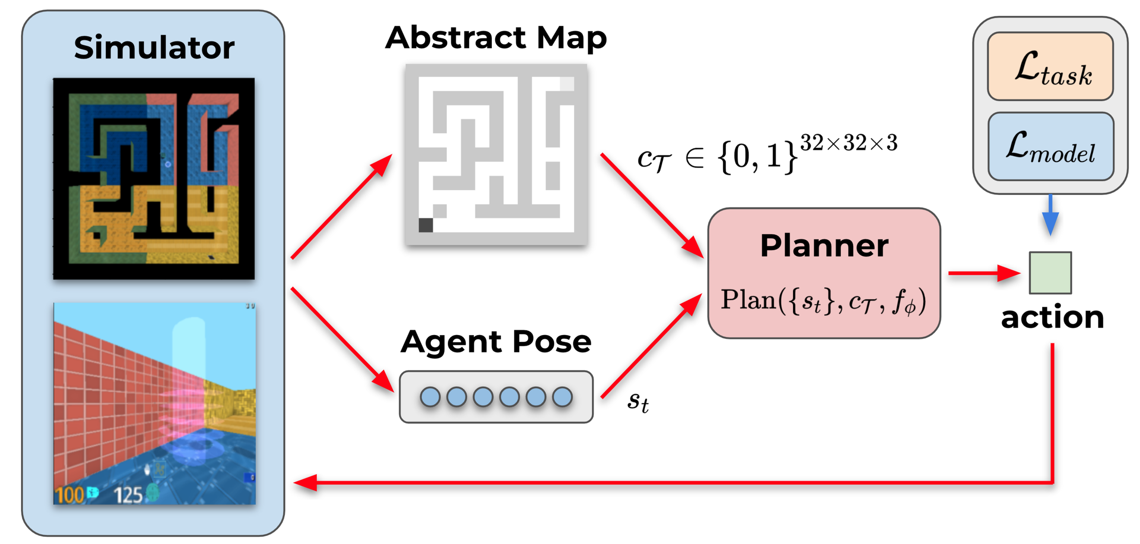
- We train the agent (planner) on a set of maps and test its generalization ability on **unseen** map layouts and goals (m, g).
- This can be viewed as a generalized goal-conditioned RL problem, while the transition function is also varying across tasks.

Formulation: Maze Navigation Environment



- The agent inputs a joint state including (1) position, (2) orientation and (3) velocity, and a task context including (1) abstract map, (2) goal and (3) start position.
- It still needs implicitly localizing its environment position to the corresponding location on the map.

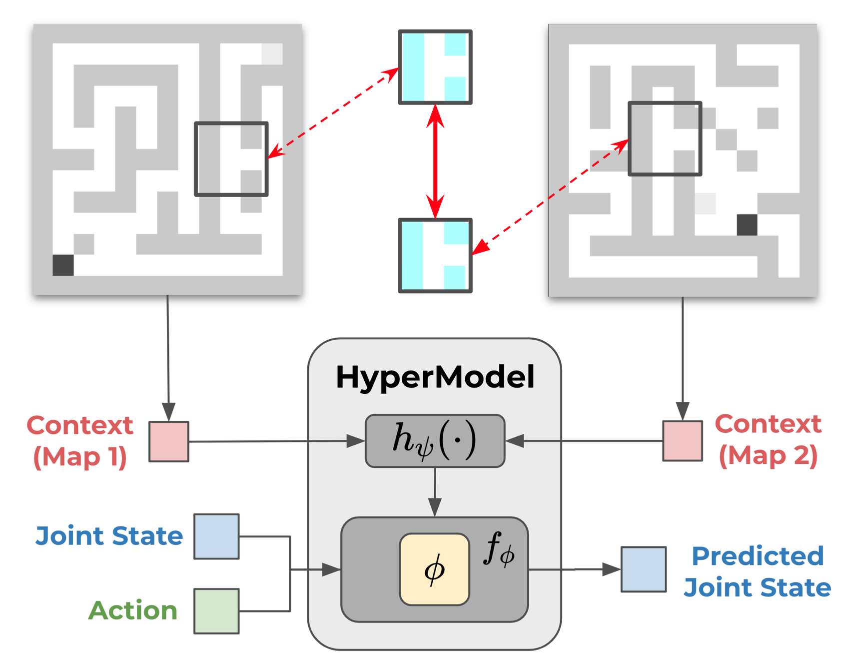
Pipeline



The overview of the pipeline. At each episode, a task is sampled by randomly selecting a map m and generating a pair of start and goal position g. The agent computes actions given context c and states. The environment simulator initializes a map and take actions.

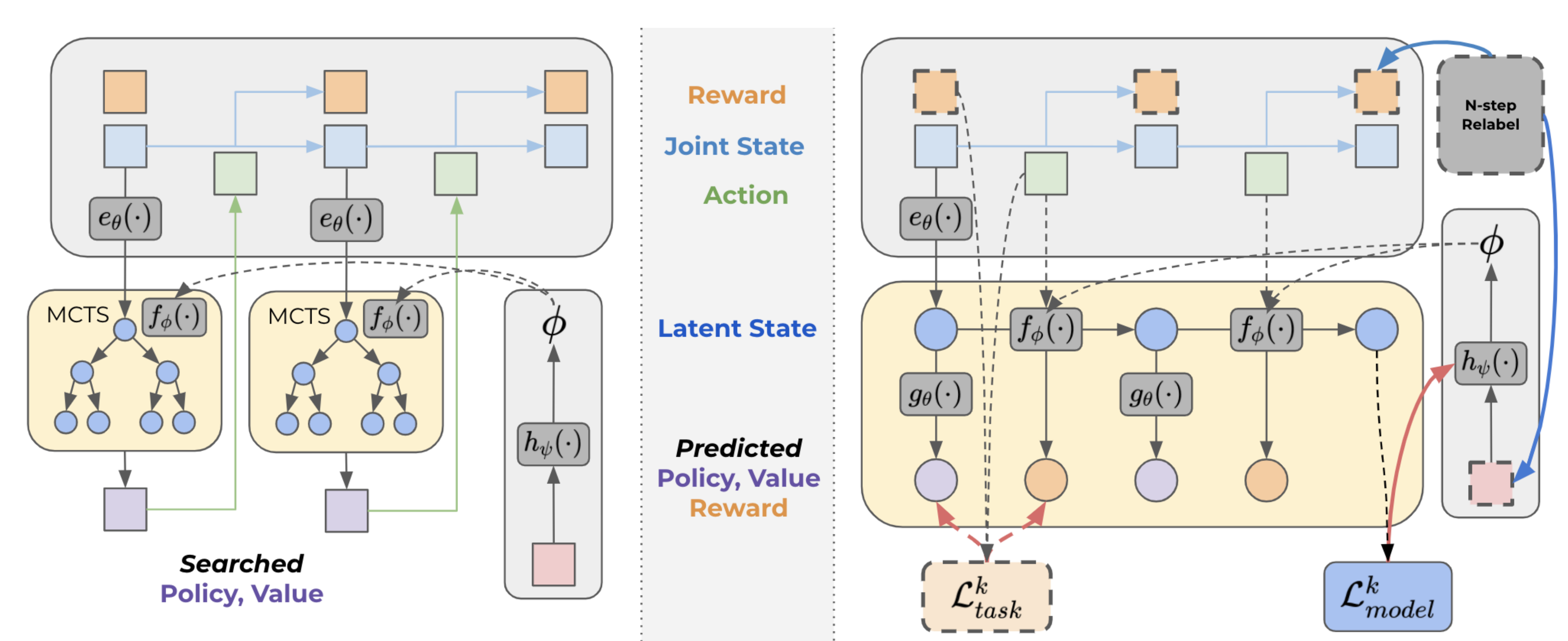
Model-based Navigation Using Abstract Maps

Task-conditioned Hypermodel



- Since the task MDP transition depends on **maps**, thus learning a reactive policy may struggle.
- We propose **hypermodel** h_ψ to *output weights of transition networks* f_ϕ for each task context c [1] and apply **model-based planning**.
- This enables **sharing computation** in abstract maps (dashed lines in left figure) and **better generalization**.

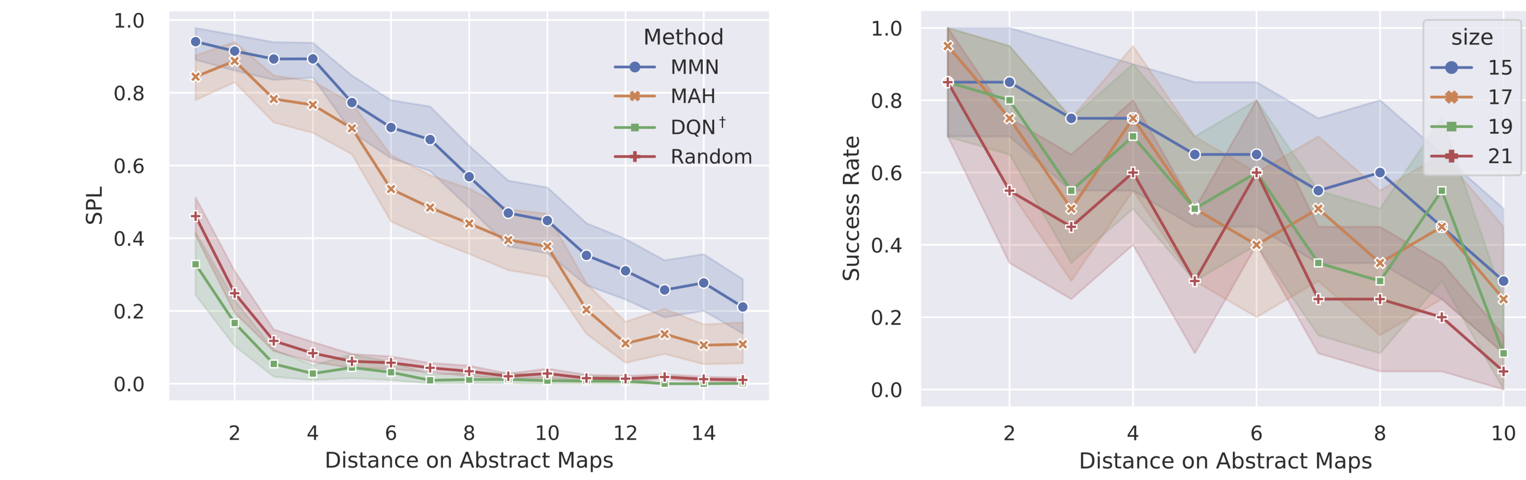
Planning with Learned Hypermodel and End-to-end Training



- We apply MCTS using the learned hypermodel to search for actions.
- In training, we use (1) **task loss** from value predictions [2] and (2) **model loss** minimizing prediction error. We propose n-**step goal relabelling** to **densify reward** [3].

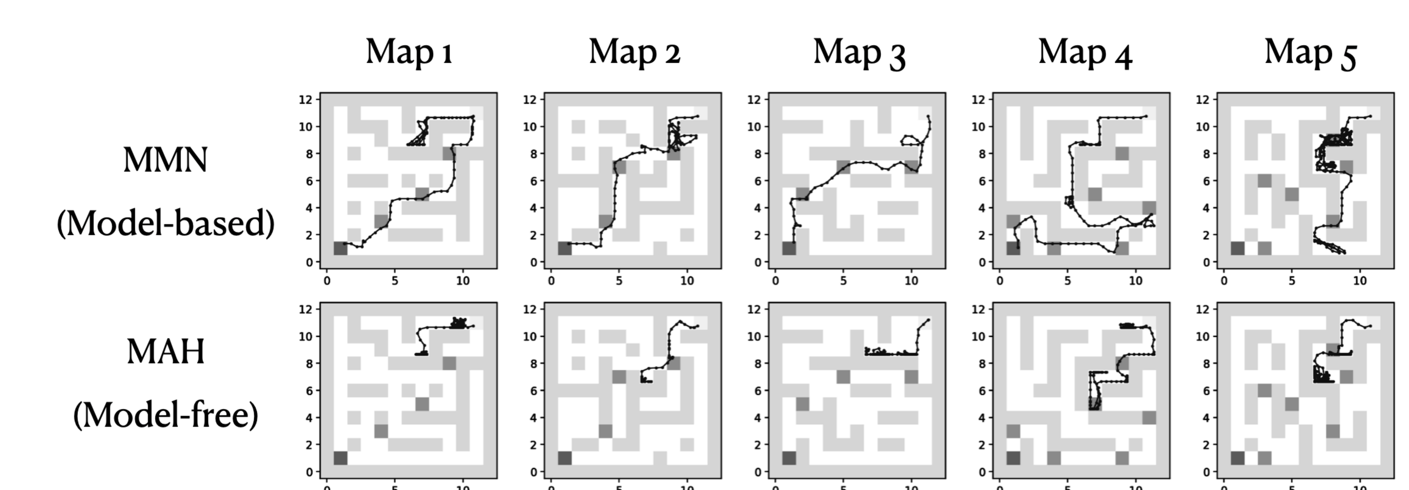
Results

Zero-shot Navigation on Novel Maps



- (Left) Zero-shot evaluation performance on 13×13 maps. Local navigation with different distances between start and goal, from 1 to 15. (Right) Performance of our method on larger maps for distance 1 to 10.
- Our hypermodel-based method (MMN) can consistently **learn faster** and **generalize better** than a strong distributed model-free baseline (MAH).

Zero-shot Hierarchical Navigation



- The sequences of subgoals are generated using a *landmark oracle*.
- MMN can **succeed on more maps** with **shorter trajectories**.

Hierarchical Navigation on Various Distances between Landmarks

Landmark Distance	1	2	3	4	5	5 (SR)
MMN	**0.61**	**0.59**	**0.68**	**0.45**	**0.63**	**0.80**
MAH	0.24	0.42	0.45	0.41	0.28	0.45
DQN†	0.00	0.00	0.00	0.00	0.00	0.00
Random	0.00	0.00	0.00	0.00	0.00	0.00

- Landmarks are provided subgoals between fixed start-goal pairs on 20 maps.
- MMN method that reads maps and uses planning outperforming others.

References

[1] David Ha, Andrew Dai, and Quoc V Le. Hypernetworks. *arXiv preprint arXiv:1609.09106*, 2016.

[2] Julian Schrittwieser, Ioannis Antonoglou, Thomas Hubert, Karen Simonyan, Laurent Sifre, Simon Schmitt, Arthur Guez, Edward Lockhart, Demis Hassabis, Thore Graepel, et al. Mastering atari, go, chess and shogi by planning with a learned model. *arXiv preprint arXiv:1911.08265*, 2019.

[3] Marcin Andrychowicz, Filip Wolski, Alex Ray, Jonas Schneider, Rachel Fong, Peter Welinder, Bob McGrew, Josh Tobin, OpenAI Pieter Abbeel, and Wojciech Zaremba. Hindsight experience replay. In *Advances in neural information processing systems*, pages 5048–5058, 2017.