Model-based Navigation in Environments with Novel Layouts Using Abstract 2-D Maps

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

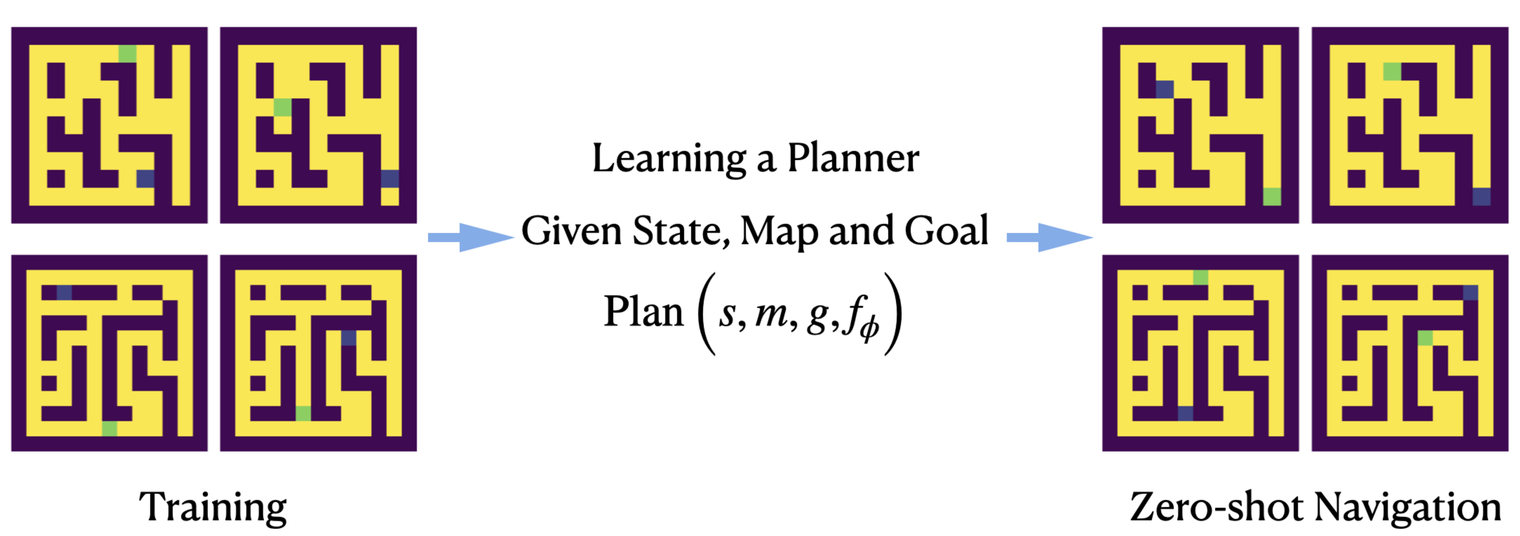
[1]Khoury College of Computer Sciences, Northeastern University

Overview

- We study **maze navigation** using **abstract 2-D maps**.
  The problem is motivated by human navigation where we are given a paper map of a novel city. The paper map can be a high-level guidance when we want navigate in a new city.
- We want to answer: How to achieve **zero-shot generalization** to novel maps and goals?
- We propose a model-based approach that learns end-to-end a hypermodel which outputs a transition network for a map and plans using the network with Monte Carlo tree search.
  It can generalize to longer distance goals on novel maps effectively by learning from limited training maps compared to model-free methods.

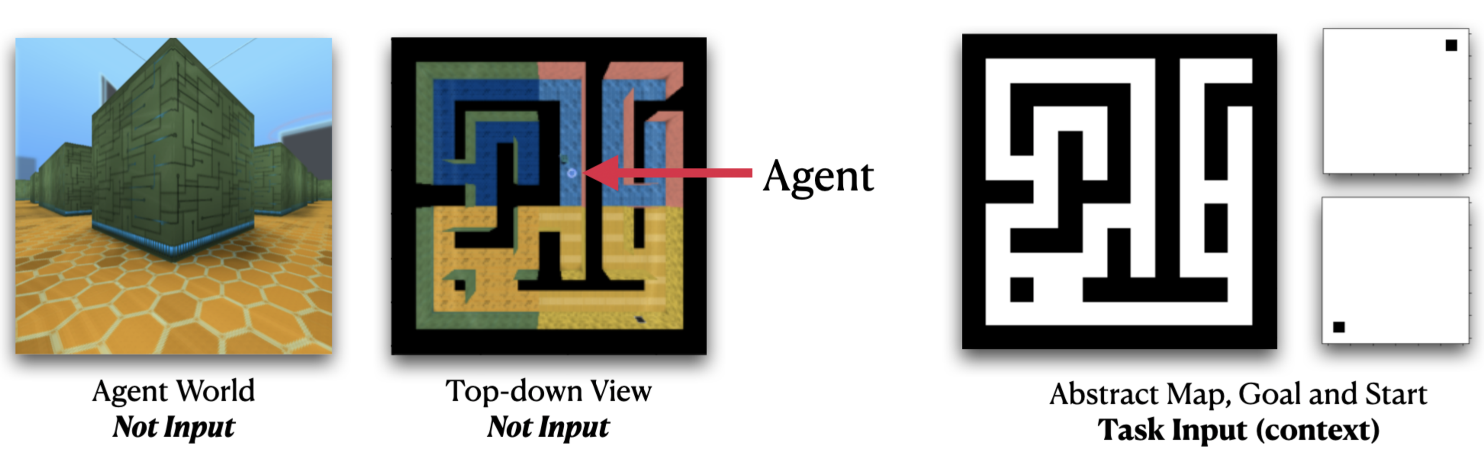
Problem Statement

Map-based *Maze* Navigation



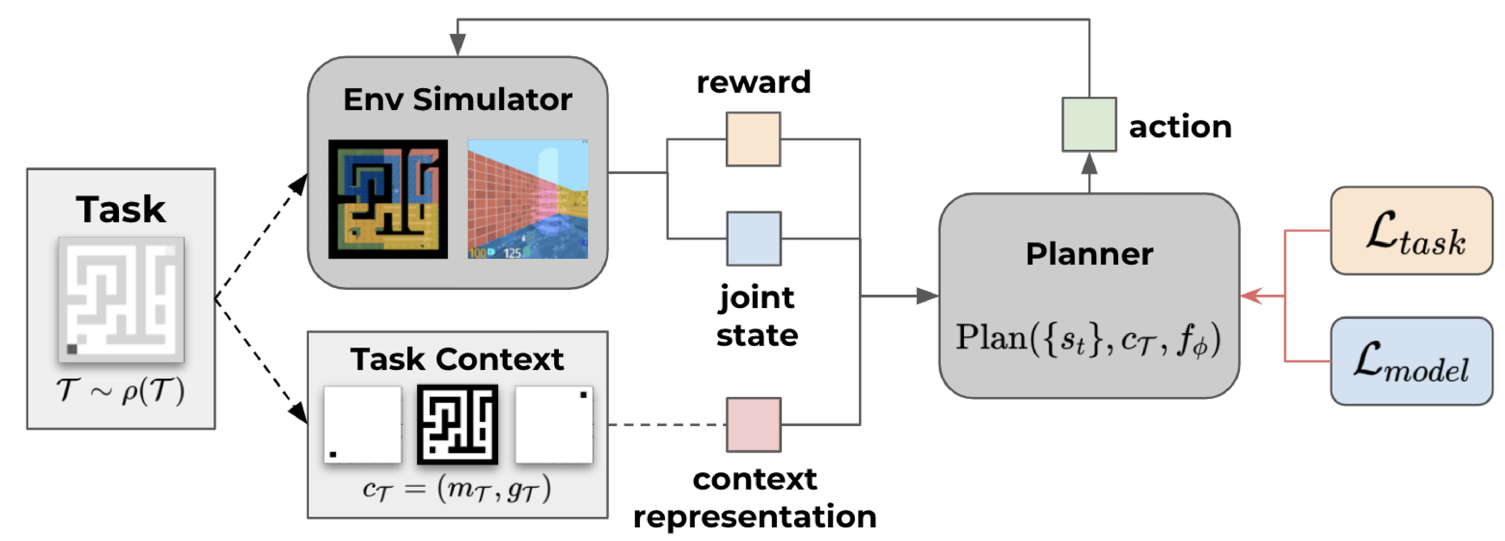
- We train the agent (planner) on a set of maps and test its generalization ability on **unseen** map layouts and goals (m, g).
- This can be viewed as a generalized goal-conditioned RL problem, while the transition function is also varying across tasks.

Maze 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.

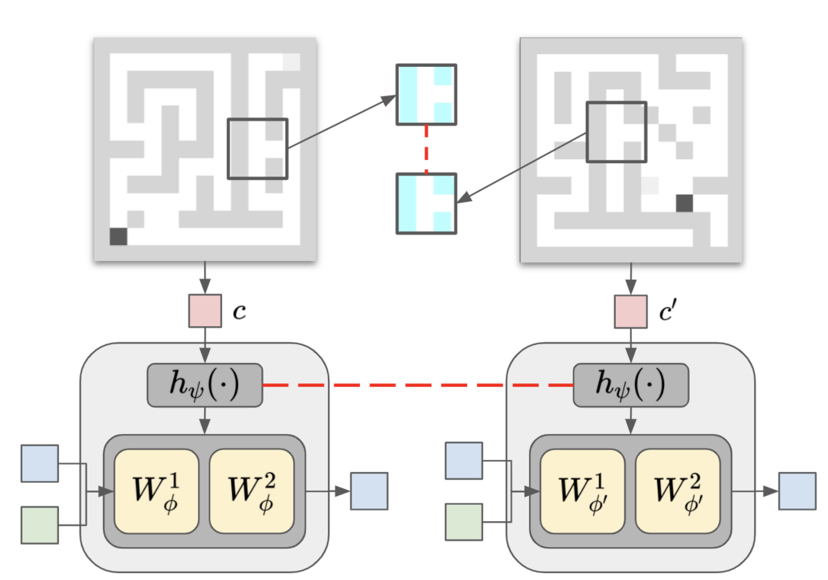
Pipeline



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.

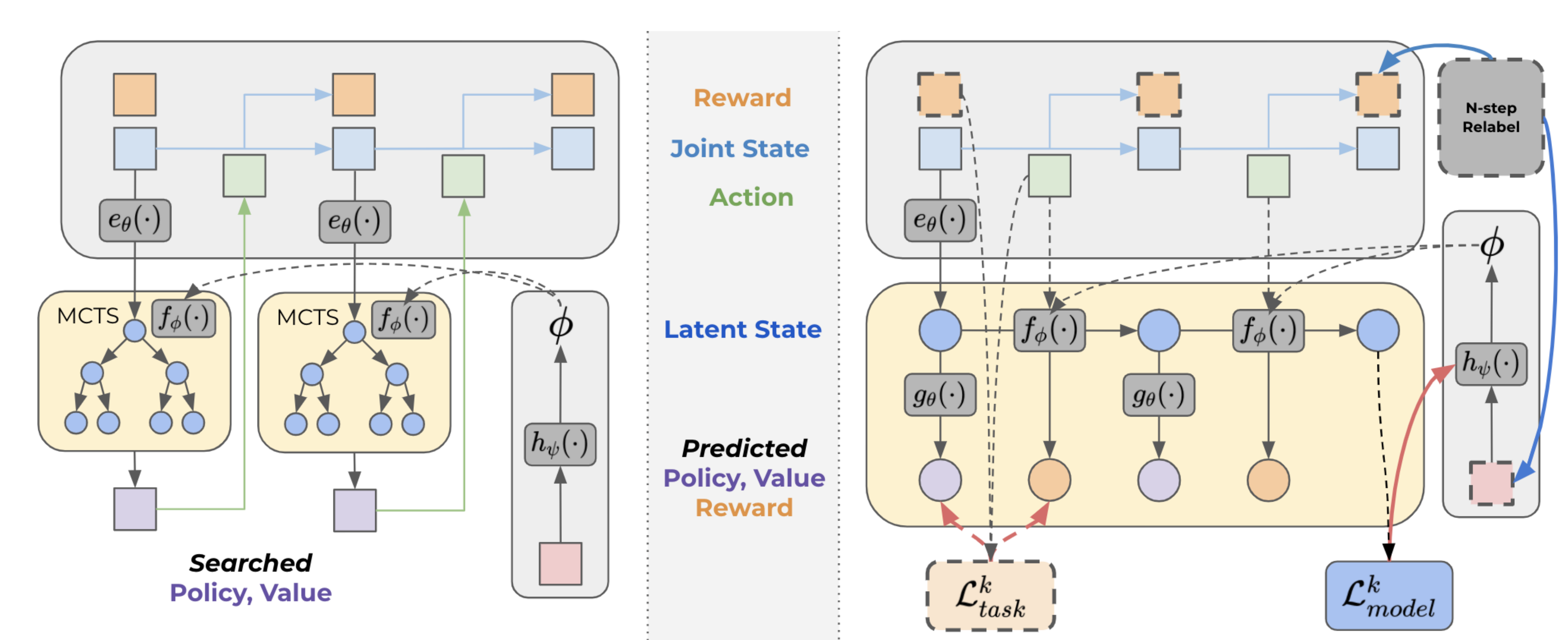
Model-based Navigation Using Abstract Maps

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**.

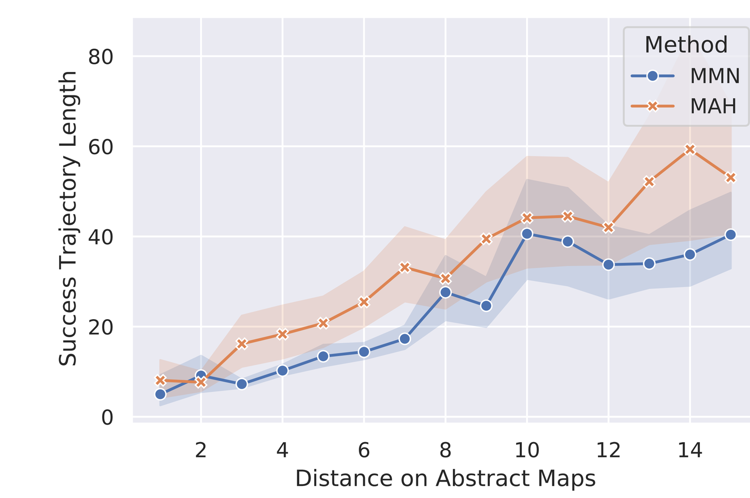
Planning with Learned Hypermodel and End-to-end Training

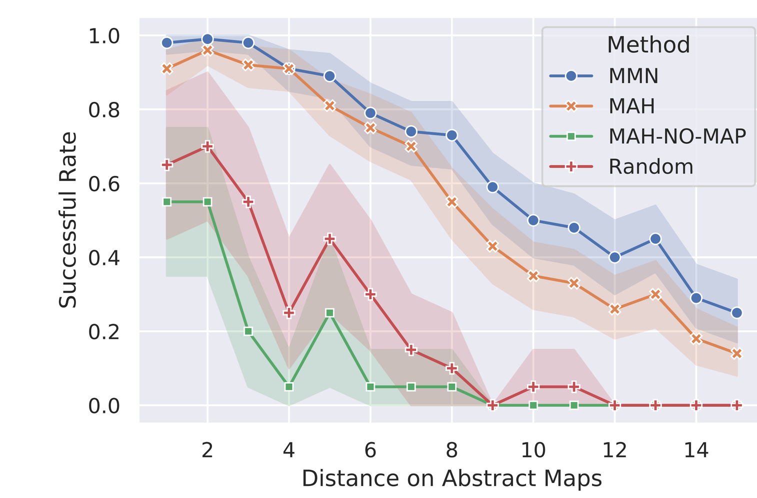


- We apply MCTS using the learned hypermodel to search for actions.
- In training, we use (1) **task loss** from value predictions [2] and (2) **model loss** minimizing prediction error. We propose n-**step goal relabelling** to **densify reward** [3].

Results

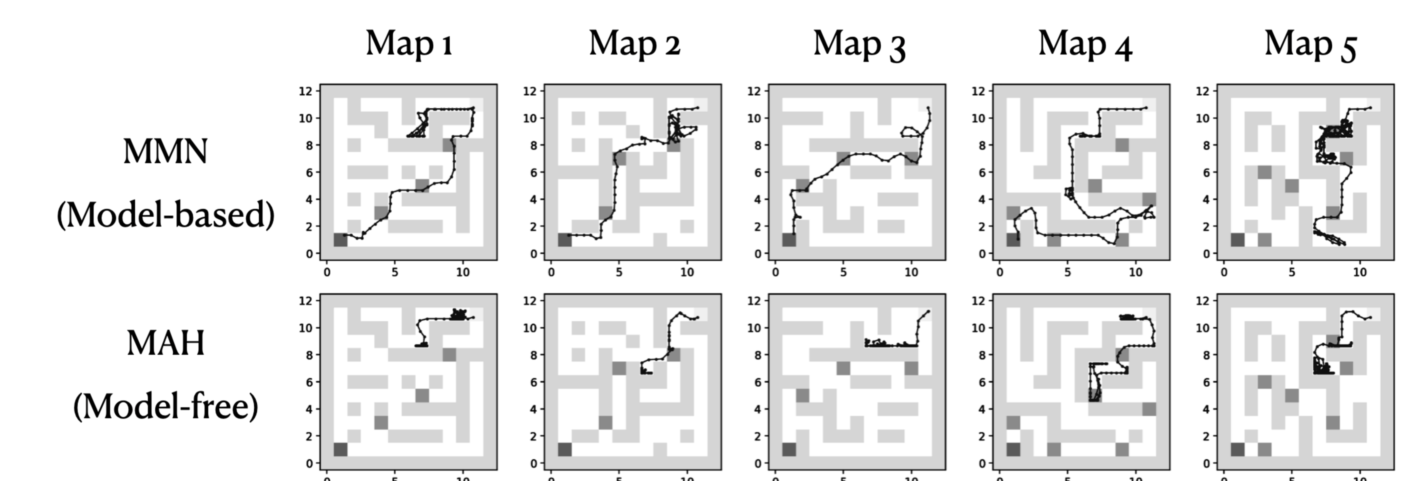
Zero-shot Navigation on Novel Maps





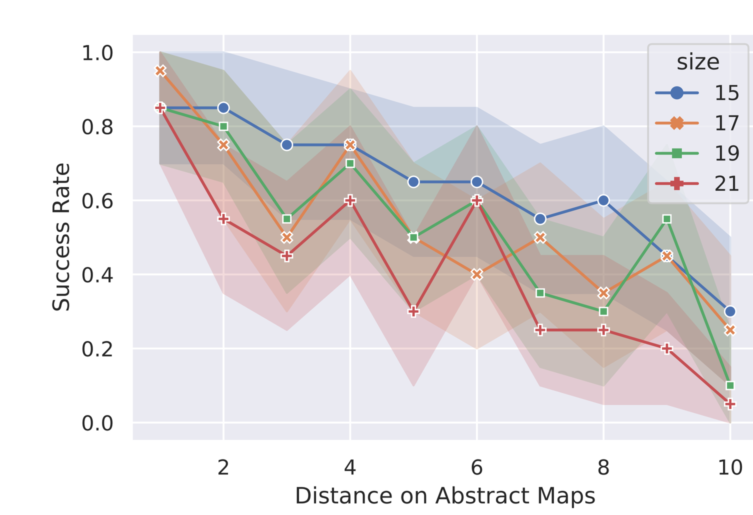
- Training on 20 13×13 maps and evaluate on 20 unseen maps.
- Our hypermodel-based method (MMN) can consistently **learn faster** and **generalize better** than a strong distributed model-free baseline (MAH).

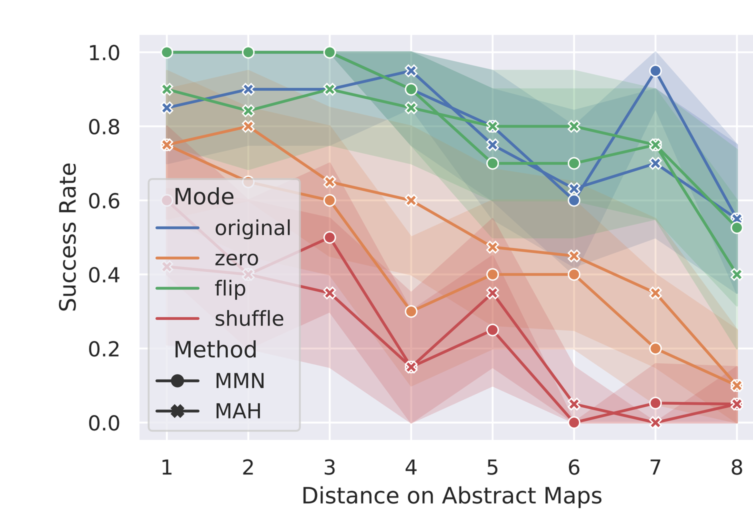
Zero-shot Hierarchical Navigation



- Use a *landmark oracle* to generate sequences of subgoals and evaluate.
- MMN can **succeed on more maps** with **shorter trajectories**.

Map Perturbation Study and Larger Map Navigation





- Zero-shot navigation on larger maps also observed successful results.
- Perturbing maps with different strategies verified its importance.

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.