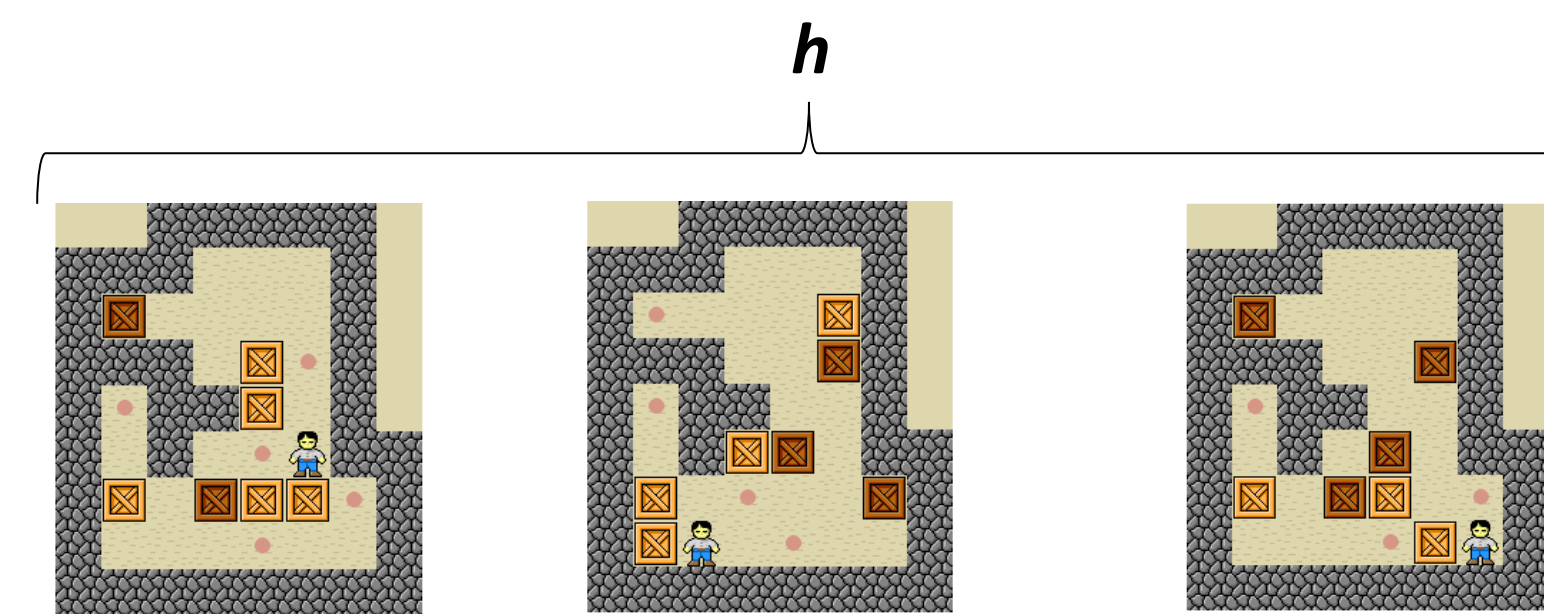


Abstract

Heuristic planning has a central role in classical planning applications and competitions. Meta-Reinforcement learning (Meta-RL) is a fast growing research area concerned with learning, from many tasks, behaviours that can quickly generalize to new tasks from the same distribution of the training ones. We make a connection between meta-reinforcement learning and heuristic planning, showing that heuristic functions meta-learned from planning problems, in a given domain, can outperform both popular domain-independent heuristics, and a heuristic learned by supervised learning.

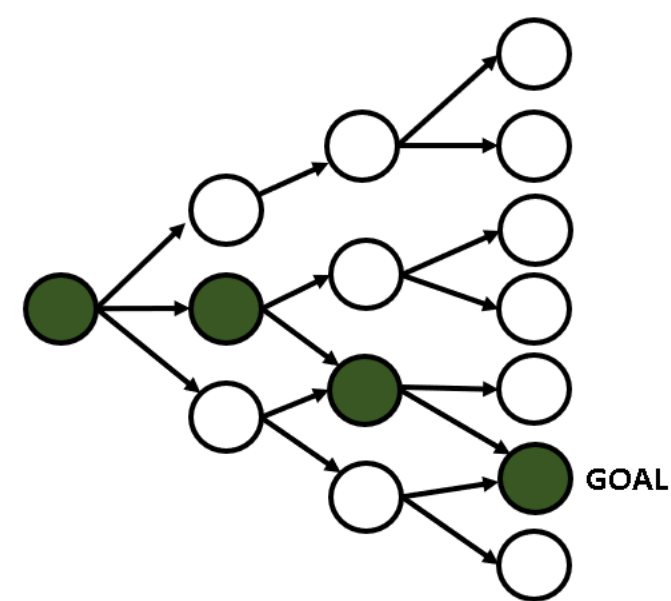
Learning heuristics with Meta-RL:

- Does not require large amounts of training tasks.
- Expands a small number of nodes to find a good solution.
- Is able to generalize to many different instances of a domain.



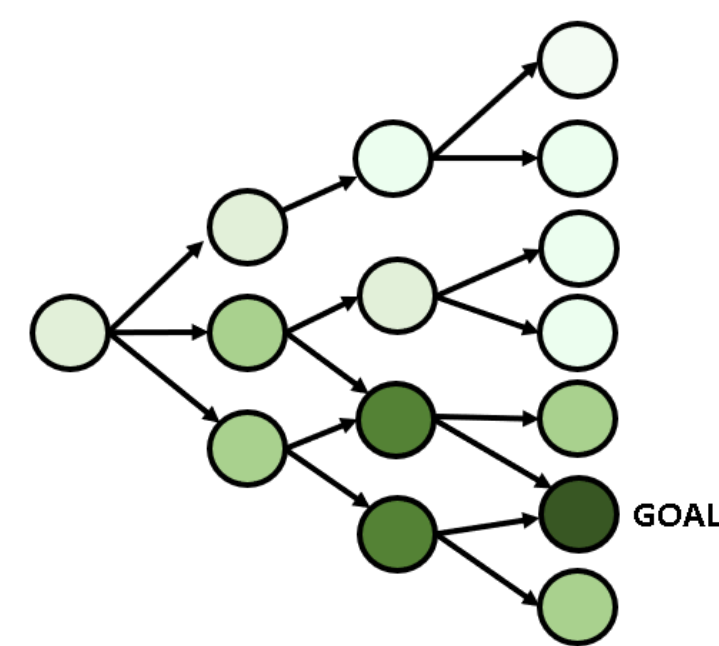
Learning Domain Dependent Heuristics

Supervised Learning



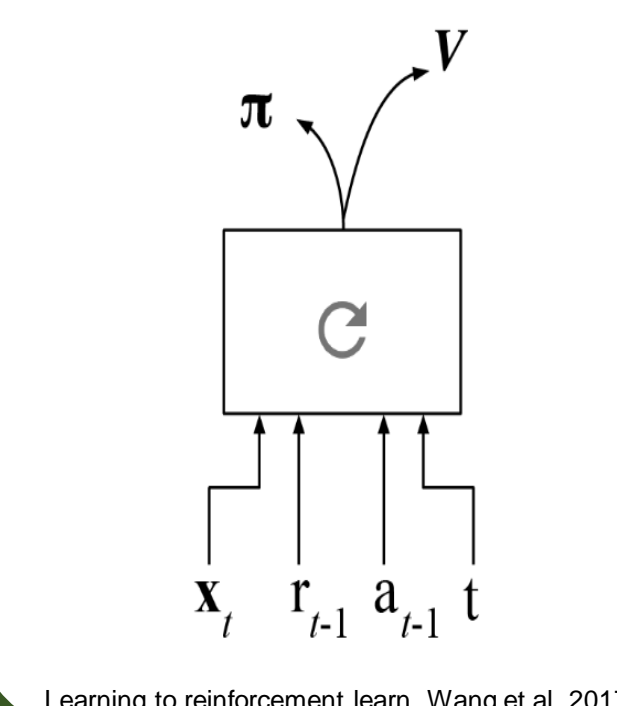
- Learns from previous solved planning problems.
- Learning is limited by the area traversed by the solved plans.
- Suboptimal states are not considered in training which can reduce generalization and harm performance of the heuristic.

Reinforcement Learning



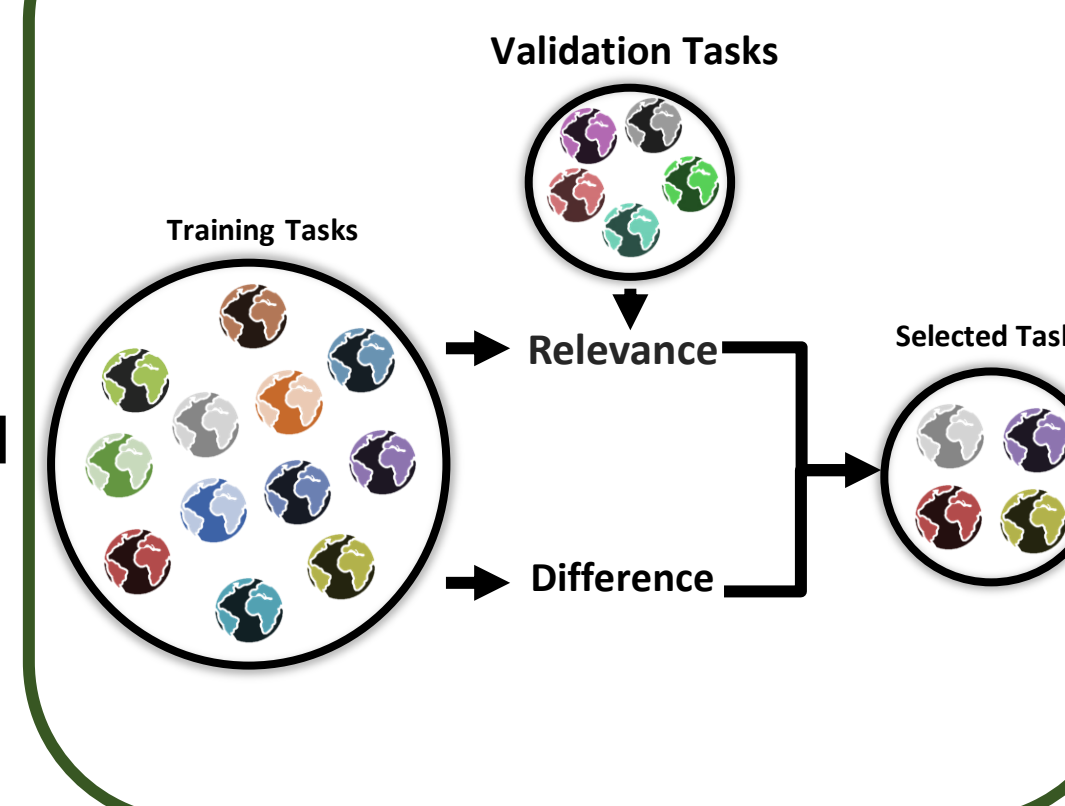
- Learns a value function for tasks modelled as Markov Decision Processes.
- Learns exploring the whole state space.
- An optimal value function is also an optimal heuristic.
- Value functions are specific to a particular task.

Context based Meta-RL



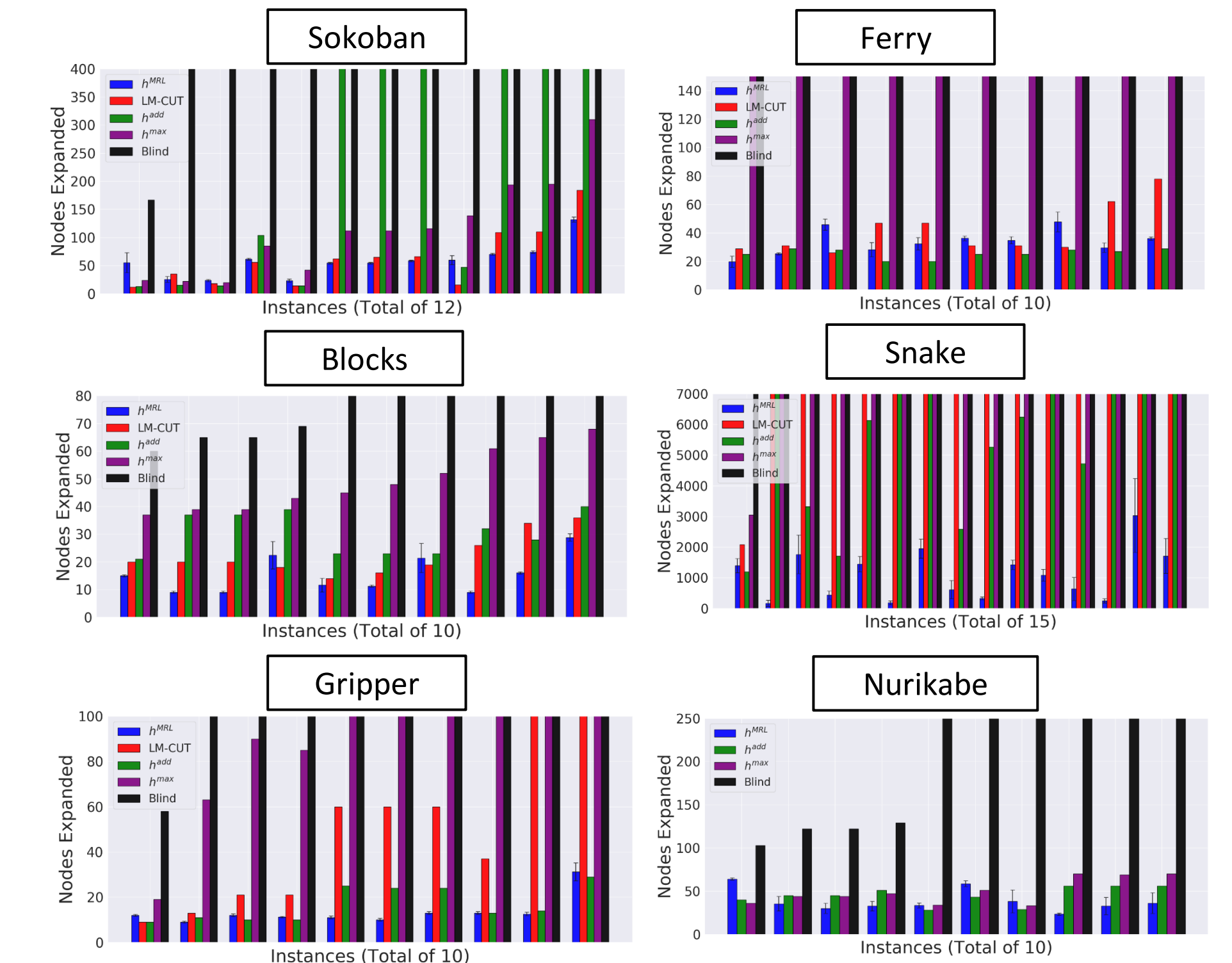
Learning to reinforcement learn, Wang et al. 2017

Task Selection



h^{MRL}

Results against Domain Independent Heuristics



Results against Supervised Learning

